ALMOST FLOW EQUIVALENCE AND THE LOOP STRUCTURE OF DIRECTED GRAPHS

by

Paulo Ventura Araujo[1]

Thesis submitted for the degree of

Doctor of Philosophy

at the Mathematics Institute,

University of Warwick.

October 1992.

[1]Partially supported by INVOTAN (Portugal).

Fátima:

Não é para a tese ficar mais gorda que esta página é só tua. É a tradição: nas outras teses fazem o mesmo, embora escrevam menos.
Um grande beijo comovido (às escondidas de quem só lê inglês).

Acknowledgements

There are two people whose contribute to the writing of this thesis was essential and whom I would like to thank most warmly.

Prof. William Parry was my supervisor since I arrived in Warwick at a moment's notice for the M.Sc. course. Besides suggesting the problems to work on, he helped me to make the most of whatever small ideas I had. His enthusiasm was an antidote to my perpetual discouragement. Thank you Bill, for having believed in me more than I did.

I would also like to thank Mike Boyle, who acted as referee when part of this thesis was submitted for publication. His reading was thourough and penetrating, and his comments and corrections were always to the point. This thesis (and particularly section 9) would be very different (and very wrong) without his help, and I was sincerely surprised and pleased with his interest.

Contents

Declaration

As far as I am aware, the work in this thesis is original except when explicitly stated otherwise.

Summary

We investigate the problem of almost flow equivalence for subshits of finite type (SFT). The problem is to decide when two suspension flows of irreducible SFT are almost everywhere one to one factors of the same suspension flow, a problem that was solved by Mike Boyle. We obtain generalizations of Boyle's result by considering the similar problems for Markov shifts and for SFT with finite group actions. We also undertake an analysis of the loop structure of directed graphs, and we reduce some problems concerning loops to the study of the loop diagram, which is a convenient form of representing the simple cycles of a given directed graph.

Introduction

We work in the classification theory of subshifts of finite type (SFT), and our starting point was to study its suspension flows (i.e., the one real-parameter flows which have global cross-sections whose Poincare map is topologically conjugate to an SFT). By the work of Bowen [B1], each basic set $(A, (\varphi_t)_{t\in\mathbf{R}})$ of an Axiom A flow is the image of some suspension flow $(X, (\sigma_t)_{t\in\mathbf{R}})$ of an SFT by a continuous map π such that $\pi\circ\sigma_t = \varphi_t\circ\pi$ for all $t \in \mathbf{R}$ and which is almost everywhere (a.e.) one to one (there exists a residual subset of A onto which π is one to one). However, we may wish to relax the condition of π commuting the flow actions and simply require that for every $x \in X$ there exists an increasing homeomorphism s: $\mathbf{R} \to \mathbf{R}$ such that $s(0) = 0$ and $\pi\circ\sigma_t(x) = \varphi_{s(t)}\circ\pi(x)$ for all $t \in \mathbf{R}$. We call these maps a.e. one to one factor maps. It is then a natural question to ask under what conditions two basic hyperbolic sets $(A_1, (\varphi_t)_{t\in\mathbf{R}})$ and $(A_2, (\psi_t)_{t\in\mathbf{R}})$ are almost flow equivalent, i.e., are both a.e. one to one factors of the same suspension flow of an SFT. The answer was provided by Mike Boyle [B3]: if none of A_1 and A_2 reduces to a single closed orbit then $(A_1, (\varphi_t)_{t\in\mathbf{R}})$ and $(A_2, (\psi_t)_{t\in\mathbf{R}})$ are almost flow equivalent.

We prove some generalizations which yield an alternative proof of this striking result. As in the original proof, we work at the level of symbolic dynamics. We base our approach on the theorem of Adler & Marcus' [AM] which states that a necessary and sufficient condition for two transitive SFT to be a.e. one to one factors of the same SFT is that they have the same topological entropy and period.

Our first generalization comes from considering suspensions of Markov shifts. Factor maps should then be both continuous and non-singular with respect to the "suspended" Markov measures. Two Markov shifts whose suspensions are both a.e. one to one factors of the same suspension of a Markov shift are called almost stochastically flow equivalent (a.s.f.e.). Our problem is then to determine necessary and sufficient conditions for two Markov shifts to be a.s.f.e.

An invariant of a.s.f.e. is the Γ-group, which is defined as the multiplicative group generated by the products of the transition probabilities along cycles. In full generality the problem of a.s.f.e. appears too difficult, and so we concentrated on the class of Markov shifts whose Γ-group is cyclic. We first prove (Proposition 5.2) that the class of suspension flows of ergodic Markov shifts thus obtained is precisely the class of suspensions of Markov shifts of maximal type (where the Markov measure considered is the measure of maximum entropy of the space supporting the measure).

As an example consider the stochastic matrices

$$P = \begin{bmatrix} \beta^{-1} & \beta^{-2} \\ 1 & 0 \end{bmatrix}, Q = \begin{bmatrix} \beta^{-2} & \beta^{-1} \\ 1 & 0 \end{bmatrix},$$

where $\beta > 0$ and $\beta^2 = 1 + \beta$, and let m_P and m_Q be the associated Markov measures on the Markov shifts Σ_P and Σ_Q. Then (Σ_P, m_P) is of maximal type and we have $h(m_P) = \log \beta > h(m_Q)$. In particular there cannot exist a measure preserving topological conjugacy $\Sigma_P \to \Sigma_Q$. Both Markov shifts have Γ-group equal to $<1/\beta>$ (the multiplicative group generated by $1/\beta$), but whereas we have, for every cycle $I_0 I_1 \ldots I_{k-1} I_0$ (with each I_j equal to either 1 or 2), the equality

$$P(I_0, I_1)\, P(I_1, I_2) \ldots P(I_{k-1}, I_0) = 1/\beta^k \qquad (0)$$

(whenever the left hand side is nonzero), such a nice condition does not hold for the matrix Q. Condition (0) is the distinctive feature of stochastic matrices of maximal type, and to prove 5.2 we show we can always modify a matrix with cyclic Γ-group to a flow equivalent matrix (i.e., a matrix which defines the same suspension space) which satisfies condition (0). For instance, Q is flow equivalent to P.

Using 5.2 and Adler-Marcus' theorem we prove our first main result (Theorem 5.1) which says that any two ergodic Markov shifts which have the same cyclic Γ-group are a.s.f.e. We obtain Boyle's theorem (stated as Theorem 6.1) by specializing 5.1 to Markov measures with Γ-group equal to $<1/2^k>$, for some $k \geq 1$.

After a deviation (sections 8 and 9) we return to the theme of almost flow equivalence in sections 10 and 11. In section 10 we consider finite group actions on SFT and its suspensions, and discuss the appropriate notion of factor map between suspension flows of SFT when a finite group action is involved. In section 11 we define almost G-flow equivalence for a finite group G and pairs (Σ_A, G) and (Σ_B, G) of SFT with fixed G-actions. Building on a theorem of Adler, Kitchen & Marcus' [AKM] which generalizes Adler & Marcus' theorem, we prove (Theorem 11.1) that if both Σ_A and Σ_B have positive entropy then (Σ_A, G) and (Σ_B, G) are almost G-flow equivalent. To prove 11.1 we show (Proposition 11.2) that for any (Σ_A, G) of positive entropy there exists (Σ_B, G) which defines the same suspension space and such that Σ_B is topologically mixing and has entropy equal to log 2.

We give a constructive proof of 11.2, and it seems to us that a proof of this result along the lines of Boyle's original proof of theorem 6.1 in [B3] is either very difficult or impossible: Boyle uses the theorem of J.Franks' [F] on the completeness of certain algebraic invariants for flow equivalence of transitive SFT, and no analogous invariants have been proved to be complete for flow equivalence of SFT with G-actions (the approach of mimicking Franks' matrix manipulations fails to work here).

The other theme of this thesis is the loop structure of directed graphs, which is the subject of section 9. The bridge between this section and the main body of the thesis is provided by section 8. The original motivation was to construct, on a given transitive SFT, a fully supported Markov measure with cyclic Γ-group having some prescribed generator β. Example 1 of section 8 illustrates a practical approach to this problem. The method consists of assigning weights which are powers of β to the simple cycles (cycles which do not cross themselves) of the associated directed graph. This assignment produces a matrix which, in case it has spectral radius one, provides the Markov measure we seek.

It becomes clear that a deeper analysis is needed, and we carry it out in section 9. We may ask, for instance, whether it is true that, for a weighted connected directed graph, the sum of the weights of the first return loops to a given vertex can be expressed in terms of the weights of the simple cycles. To answer this question we first show (Proposition 9.2) that each loop (based at a given vertex) has a factorization where the factors are simple cycles (in some order), and that sections 10 and 11. In section 10 we consider finite group actions on SFT and its suspensions, and discuss the appropriate notion of factor map between suspension flows of SFT when a finite group action is involved. In section 11 we define almost G-flow equivalence for a finite group G and pairs (Σ_A, G) and (Σ_B, G) of SFT with fixed G-actions. Building on a theorem of Adler, Kitchen & Marcus' [AKM] which generalizes Adler & Marcus' theorem, we prove (Theorem 11.1) that if both Σ_A and Σ_B have positive entropy then (Σ_A, G) and (Σ_B, G) are almost G-flow equivalent. To prove 11.1 we show (Proposition 11.2) that for any (Σ_A, G) of positive entropy there exists (Σ_B, G) which defines the same suspension space and such that Σ_B is topologically mixing and has entropy equal to log 2.

We give a constructive proof of 11.2, and it seems to us that a proof of this result along the lines of Boyle's original proof of theorem 6.1 in [B3] is either very difficult or impossible: Boyle uses the theorem of J.Franks' [F] on the completeness of certain algebraic invariants for flow equivalence of transitive SFT, and no analogous invariants have been proved to be complete for flow equivalence of SFT with G-actions (the approach of mimicking Franks' matrix manipulations fails to work here).

The other theme of this thesis is the loop structure of directed graphs, which is the subject of section 9. The bridge between this section and the main body of the thesis is provided by section 8. The original motivation was to construct, on a given transitive SFT, a fully supported Markov measure with cyclic Γ-group having some prescribed generator β. Example 1 of section 8 illustrates a practical approach to this problem. The method consists of assigning weights which are powers of β to the simple cycles (cycles which do not cross themselves) of the associated directed graph. This assignment produces a matrix which, in case it has spectral radius one, provides the Markov measure we seek.

It becomes clear that a deeper analysis is needed, and we carry it out in section 9. We may ask, for instance, whether it is true that, for a weighted connected directed graph, the sum of the weights of the first return loops to a given vertex can be expressed in terms of the weights of the simple cycles. To answer this question we first show (Proposition 9.2) that each loop (based at a given vertex) has a factorisation where the factors are simple cycles (in some order), and that any loop is determined by its factorization and its base vertex. This allows us to reduce all questions having to do with loops to the study of a certain undirected graph which we call the loop diagram. We are then able to deduce the formula for the sum we consider, and in the course of the proof we deduce other similar formulas. The two theorems of this section (Theorems 9.1 and 9.6) address the question of how to know whether a given assignment of weights to the simple cycles of a directed graph produces a Markov measure such that the assigned weights are the products of the transition probabilities along the cycles.

We feel that much of the interest of section 9 lies in the methods we introduce, avoiding calculations with matrices and working instead directly with graphs.

In the final remarks (section 12) we point out an interesting connection between Williams' conjecture and the problem of deciding whether some known invariants for stochastic flow equivalence are complete. We show that a positive answer to this question, even in the simplest case of matrices with cyclic Γ-group, would also answer Williams' conjecture for irreducible integer matrices affirmatively.

Sections 1 to 9, together with an appendix, make up an article which has been recommended for publication in Ergodic Theory & Dynamical Systems, and we reproduce it here with its title and in the form it was submitted for publication.

A stochastic analogue of a theorem of Boyle's on almost flow equivalence

by

Paulo Ventura Araujo[1]

Centro de Matematica, Faculdade de Ciencias, 4000 Porto, Portugal.

Abstract. We study a new topological classification of suspension flows on subshifts of finite type, and obtain a new proof of a theorem of Boyle's which states that, in an appropriate sense, all such flows are alike. We prove that the stochastic version of this classification is non-trivial by exhibiting a certain invariant, and show that this invariant is complete in a particular case, although not in general. Symbolic flows are important as models of basic sets of Axiom A flows, and so we discuss the significance of our results for this latter type of flow.

1. *Introduction.*

Notwithstanding steady progress, the classification of subshifts of finite type (SFT) is not as yet complete. For example, we still don't have a satisfactory answer as to when two irreducible SFT are topologically conjugate, or when one is a factor of the other. Partly because these are indeed difficult questions, and partly to understand better the role of certain invariants and certain types of endomorphisms, mathematicians were lead to consider variants on these problems and different types of classification. The type of classification we're concerned with here (almost flow equivalence) is the flow analogue of Adler & Marcus' almost topological conjugacy [AM]. Boyle's surprising result is that any two non-trivial suspended flows of irreducible SFT are almost flow equivalent (theorem 6.1). In contrast, we find that the

[1]Financially supported by INVOTAN (Portugal)

introduction of Markov measures produces a wealth of equivalence classes for the stronger notion of almost stochastic flow equivalence (theorem 5.1).

According to the Adler & Marcus' definition, two SFT are almost conjugate if they are factors (i.e. homomorphic images) of the same SFT by maps which are almost conjugacies, in a precise sense. Similarly, we say that two SFT are almost flow equivalent if their suspension flows are almost homeomorphic images of the same flow. The main theorem then reads as follows:

6.1. Theorem (Boyle [B3]). *Any two irreducible SFT with positive topological entropy are almost flow equivalent.*

To give a new proof of 6.1 we first prove a stochastic version of it:

5.1. Theorem. *The Γ-group is an invariant of almost stochastic flow equivalence. If two irreducible Markov shifts have the same cyclic Γ-group then they are almost stochastically flow equivalent.*

The main difference between the present proof and Boyle's proof of 6.1 is that the latter uses Krieger's embedding theorem and Franks' invariants for flow equivalence, whereas we rely essentially on the Adler & Marcus' theorem (which we here state as theorem 2.2). Theorem 6.1 translates to basic sets of hyperbolic flows as follows:

7.2. Theorem (Boyle [B3]). *Let Λ_1 and Λ_2 be hyperbolic basic sets of the flows φ_t and ψ_t. If neither of the flows $\varphi_t|_{\Lambda_1}$ and $\psi_t|_{\Lambda_2}$ consists of a single periodic orbit then they are almost flow equivalent.*

Section 2 contains basic definitions and results and settles the notation. Section 3 is devoted to introductory material on flows. In contrast with the situation for the conjugacy of

SFT, the problem of topological equivalence of suspension flows of irreducible SFT is now entirely solved: Franks [F] proved that two easily computable algebraic invariants completely classify flow equivalence classes. We include a proof of the stochastic version of the main result of [PSu], which provided a first characterization of these classes. This is our corollary 3.2.

We explain what the Γ-group is in section 4. We remark that there are simple examples of irreducible Markov shifts which have cyclic Γ-group with generator an arbitrary Perron number (see [L] for the definition of Perron number), and so Theorem 5.1 implies that almost stochastic flow equivalence is far from trivial.

Section 5 is devoted to the proof of 5.1, and in section 6 we derive theorem 6.1 from 5.1. To do this, we construct, on any given irreducible SFT, a Markov measure whose Γ-group is generated by a power of 1/2.

In section 7 we comment on the implications of theorem 6.1 for basic sets of Axiom A flows, and particularly for the so-called Smale flows.

Section 8 contains an example and a counter-example. The example illustrates an interesting practical method of constructing, on a given irreducible SFT, a Markov measure with specified Γ-group; this is based on the so-called loop polynomial. The counter-example shows that the Γ-group is not a complete invariant of almost stochastic flow equivalence.

Section 9 sets the foundations of the method we introduce in example 8.1. It contains a nice formula for the sum of the weights of the first return loops to a given vertex in a directed graph. For its proof we introduce the loop diagram, which describes the loop structure of a given directed graph in a very convenient form.

In the appendix we include a generalization of proposition 5.2 proving that within each non-trivial stochastic flow equivalence class there exists a Markov shift where the quotient group Γ/Δ is infinite cyclic.

Much of this work was done at the Mathematics Institute, University of Warwick, England. Bill Parry suggested the problems and endured my idiosyncrasies with the utmost patience: my deeply-felt thanks to him. Thanks to Mark Pollicott for reading and commenting earlier versions of this article. And thanks to Mike Boyle for his suggestions and his acute

comments and corrections, which improved this article immeasurably.

2. *Generalities.*

Let A be a non-negative $k \times k$ matrix. A is *irreducible* if for any $1 \leq I, J \leq k$ there exists $n > 0$ such that $A^n(I,J) > 0$. A is *aperiodic* if the same n can be chosen for all I, J. The *period* of A is the g.c.f. of the integers n such that $A^n(I,I) > 0$ (for any $1 \leq I \leq k$ - this is independent of I). A is therefore aperiodic if and only if its period is 1.

Given an irreducible integer matrix A we have a directed graph (again called A) naturally associated with A: its *vertices* are numbered 1, 2,..., k and there are A(I,J) *edges* from I to J; V_A denotes the set of vertices, and E_A the set of edges. A *cycle* or *loop* in A is a finite closed path. A *simple cycle* is a cycle which goes only once through each of its vertices. The *length* $\ell(u)$ of the path u is the number of *edges* in it.

The *subshift of finite type* (*SFT*) Σ_A is the set of all doubly infinite paths in the graph A. More formally $\Sigma_A = \{(x_n)_{n \in \mathbb{Z}} \in E_A^{\mathbb{Z}}: f(x_n) = i(x_{n+1}) \ \forall\ n \in \mathbb{Z}\}$ (where f(x) and i(x) are respectively the *final* and *initial* vertices of the edge x). Equipped with the product topology (of the discrete topology in E_A), Σ_A is a totally disconnected compact topological space. A basis for its topology is formed by the *cylinder sets* ${}_j[e_0 \ldots e_s] = \{x \in \Sigma_A : x_j \ldots x_{j+s} = e_0 \ldots e_s\}$. If $I \in V_A$ then we denote ${}_j[I] = \{x \in \Sigma_A: i(x_j) = I\}$. The *shift* map σ_A (or just σ where no confusion is possible) is a homeomorphism of this space, defined by moving sequences one position to the left, i.e. $\sigma(x_n)_{n \in \mathbb{Z}} = (x_{n+1})_{n \in \mathbb{Z}}$. The *topological entropy* of the system (Σ_A, σ_A) is the natural logarithm of the maximum eigenvalue of the matrix A, which is a well-defined positive number, since A is non-negative and irreducible.

An *almost conjugacy* between SFT is a *factor map* (continuous, shift-commuting, surjective) which is also a.e. 1 to 1. This definition is independent of the specific choice of a (supported) measure, as our next proposition makes clear. A good reference for the proof is [AM].

2.1. Proposition. *Let A and B be irreducible integer matrices and $\pi: \Sigma_A \to \Sigma_B$ a factor map. Then the function π is finite to one if and only if σ_A and σ_B have the same topological entropy. If π is finite to one then the following statements are equivalent:*

(i) *π is a.e. one to one with respect to some fully supported ergodic measure on Σ_A;*

(ii) *π is a.e. one to one with respect to all fully supported ergodic measures on Σ_A;*

(iii) *there is a point in Σ_B which has a single pre-image;*

(iv) *every doubly transitive point in Σ_B (a point whose forward and backward orbits are both dense) has a single pre-image;*

(v) *there exists a resolving block for π, i.e. there exists a path $e_0 \ldots e_s$, a number $0 \leq j \leq s$, and a vertex $J \in V_A$ such that, for any $y \in \pi^{-1}({}_0[e_0 \ldots e_s])$, $\pi(y_j) = J$.*

Two SFT are *almost topologically conjugate* (or, abreviately, *almost conjugate* - but the reader is warned that saying two SFT are almost conjugate is not saying there exists an almost conjugacy from one onto the other) if they are a.e. one to one factors of the same SFT. This concept was introduced by Adler and Marcus, who found it to be the natural topological counterpart of the measure-theoretical classification of Markov shifts. Measure-theoretical entropy and period classify irreducible Markov shifts for the (metric) isomorphism relation, whereas topological entropy and period classify irreducible SFT for the relation of almost (topological) conjugacy.

2.2. Theorem (Adler and Marcus [AM]). *Two irreducible SFT are almost conjugate if and only if they have the same topological entropy and period.*

If P is a stochastic matrix, and p is the probability vector such that pP=p, the *Markov measure* μ_P is defined on cylinders by $\mu_P({}_j[I_0 I_1 \ldots I_s]) = p(I_0)P(I_0,I_1) \ldots P(I_{s-1},I_s)$. The entries of P are called *transition probabilities*. If P^0 is the 0-1 matrix which results from substituting 1 for each non-zero entry of P, then the corresponding SFT Σ_{P0} is the support of μ_P.

Now let A be a 0-1 irreducible matrix, λ its maximum eigenvalue, v the corresponding column (right) eigenvector, and P the stochastic matrix defined by $P(I,J) = A(I,J)v(J)/\lambda v(I)$. Amongst the many possible Markov measures on Σ_A, μ_P is particularly important:

2.3. Proposition (Parry [P]). *μ_P is the unique invariant measure on Σ_A with maximum entropy, equal to $\log \lambda$.*

These measures μ_P are called *maximal measures.*

We have in the preceding discussion considered only 0-1 matrices, but made no such restriction on the definition of SFT. The problem is that transition probabilities along different parallel edges should be allowed to be different, and yet we'd like to record these probabilities in matrix form. This leads to the consideration of exponential functions, an approach outlined in [PT1] and developed in [MT].

Let then A be an irreducible non-negative integer matrix. We now assign *weights* to the edges of A, i.e. we consider a function $wt: E_A \longrightarrow \mathbb{R}^+$; to record these weights in matrix form we use the semi-ring $\mathbb{Z}^+(\exp)$ (exp is the set $\{t \to a^t : a > 0\}$ of exponential functions), and define the matrix A_{wt} by $A_{wt}(I,J)(t) = \sum_{i(e)=I,\ f(e)=J} wt(e)^t$.

For brevity we call *exponential matrices* to those matrices whose entries are in $\mathbb{Z}^+(\exp)$. Two exponential matrices $A(t)$, $B(t)$ are *cometric* (resp. *conjugate*) if there are a number $\lambda > 0$ and an invertible non-negative diagonal matrix D such that $B(t) = \lambda^t D^t A(t) D^{-t}$ (resp. $B(t) = D^t A(t) D^{-t}$).

If the matrix $A_{wt}(1)$ is stochastic then it defines a Markov measure $m_{A_{wt}}$ on the SFT Σ_A, for which the transition probabilities on the edges between I and J are given by the basis of the exponential functions in the sum $A_{wt}(I,J)$. Otherwise there is a unique exponential matrix B such that B is cometric with A_{wt} and $B(1)$ is stochastic, and the measure $m_{A_{wt}}$ is then the measure defined by B. [Hence the terminology, for two cometric matrices define the same Markov measure.]

We can also start with an exponential matrix A, which will define a Markov measure supported on the space $\Sigma_{A(0)}$. The *weight* of the path $u = e_1 \ldots e_n$ is then defined by $wt_A(u) = wt_A(e_1) \ldots wt_A(e_n)$.

2.4. Proposition [MT]. *If A and B are irreducible exponential matrices, then a necessary and sufficient condition for a finite to one, shift-commuting, continuous map $\pi\colon \Sigma_A \longrightarrow \Sigma_B$ to be measure-preserving (m.p.) is that for some constant $\alpha > 0$ the equality*

$$wt_A(u) = \alpha^{\ell(u)}\, wt_B(\pi(u)) \tag{1}$$

holds for all cycles u in A, where $\pi(u)$ is the unique cycle in B such that $\ell(\pi(u)) = \ell(u)$ and $\pi(u)^\infty = \pi(u^\infty)$.

The number α above is the quotient $\beta(B)/\beta(A)$ of the spectral radii of B(1) and A(1). This proposition and the fact that the weights on cycles are unchanged allow us to disregard whether a given exponential matrix A with $\beta(A)=1$ is such that A(1) is actually stochastic - and, unless otherwise stated, by irreducible exponential matrix we shall mean a matrix whose value at the point 1 has spectral radius 1.

Thus we are interested in maps which are continuous and preserve measures. A *block isomorphism* is a measure-preserving topological conjugacy between Markov shifts; and the analogue of Adler & Marcus' almost conjugacy is *almost block isomorphism*, where the additional condition is that maps should be measure-preserving.

3. *Flows.*

Let A be an irreducible integer matrix and $f\colon \Sigma_A \rightarrow \mathbb{R}^+$ a continuous function. The *suspension space* Σ_A^f is defined as the quotient space

$$\Sigma_A^f = \{(x,t) \in \Sigma_A \times \mathbb{R}: 0 \leq t \leq f(x)\}/\sim,$$

where $\sim$ is the equivalence relation that identifies $(x,f(x))$ and $(\sigma x,0)$ for each x. The *flow* $(\sigma_A^s)_{s\in\mathbb{R}}$ (or simply σ^s) moves points vertically on Σ_A^f at unit speed. Thus $\sigma^s[x,t] = [x,s+t]$ for $t<f(x)$ and small s, and for other values of s we use the identifications.

When A is an exponential matrix we define the probability measure m_A^f by the formula

$$\int_{\Sigma_A^f} F\,dm_A^f = \int_{\Sigma_A} \left(\int_0^{f(x)} F[x,t]\,dt\right) dm_A(x) \Big/ \int_{\Sigma_A} f\,dm_A, \text{ for } F \in C(\Sigma_A^f). \qquad (2)$$

With respect to the flow, the measure m_A^f is invariant and ergodic. By a result in [A], all $(\sigma_A^s)_s$-invariant measures have the form μ^f, for some shift-invariant measure μ. Also μ^f is ergodic (with respect to the flow) if and only if μ is ergodic (with respect to the shift).

Any two suspensions of the same base space are essentially the same, in that a simple continuous reparametrization of the orbits turns one flow into the other. For this reason we state our definitions and results using the standard suspension, i.e. the suspension with constant first return time of one unit.

A *factor map* (for suspensions of Markov shifts) is a map $\pi: \Sigma_A^1 \to \Sigma_B^1$ with the following properties:

(i) π is continuous;

(ii) π is surjective;

(iii) π is non-singular (the measure $\pi^* m_A^1$ is equivalent to m_B^1);

(iv) π sends orbits onto orbits, preserving the orientation;

(v) the restriction of π to any orbit is a local homeomorphism onto its image.

The concept of factor map for suspensions of SFT is the same, except that we drop condition (iii). A *topological* (resp. *stochastic*) *flow equivalence* is a factor map between suspensions of SFT (resp. Markov shifts) which is also injective and therefore a conjugacy. Parry & Sullivan [PSu] introduced and studied the concept of flow equivalence, showing that two irreducible integer matrices A and B are flow equivalent (i.e. define flow equivalent suspension spaces) if and only if there is a sequence of integer matrices

$$A = A_1, A_2, \dots, A_r = B$$

such that for each $0 < i < r$ either A_i and A_{i+1} define conjugate subshifts or one of them is the matrix

$$\begin{bmatrix} a_{11} & \cdots & a_{1n} \\ & \cdots & \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

whereas the other is

$$\begin{bmatrix} 0 & a_{11} & & a_{1n} \\ 1 & 0 & \cdots & 0 \\ 0 & a_{1n} & & a_{2n} \\ & & \cdots & \\ 0 & a_{n1} & & a_{nn} \end{bmatrix}.$$

The process of passing from the former to the latter matrix is called *expansion*.

In the sequel, following [PT2], we present a proof of this result in the more general setting of stochastic flow equivalence. We state it as corollary 3.2.

Again consider an irreducible exponential matrix A, and let a function $k: V_A \to \mathbb{N}$ be given. The matrix A_k is defined as follows: the vertices are (I,i) ($I \in V_A$, $1 \le i \le k(I)$); the non-zero entries are $A_k((I,i),(I,i+1)) = 1$ (if $1 \le i < k(I)$) and $A_k((I,k(I)),(J,1)) = A(I,J)$ (if $A(I,J) \neq 0$). Clearly, A_k is obtained from A by a number of expansions.

Given a factor map $\pi: \Sigma_A^1 \to \Sigma_B^1$, the equations $\pi[x,s] = \sigma_B^{h_x(s)}\pi[x,0]$, $h_x(0) = 0$ ($x \in \Sigma_A$) define a unique continuous function h_x, which is a homeomorphism of the real line. It gives the distance $h_x(s)$ the image point of $[x,0]$ has moved when the point $[x,0]$ has travelled a distance s. Letting $h(x)=h_x(1)$, the composite function

$$\Sigma_A^h \longrightarrow \Sigma_A^1 \xrightarrow{\pi} \Sigma_B^1$$

$$[x,t] \to [x,h_x^{-1}(t)]$$

is a semiconjugacy (i.e., it commutes with the flow), and it is also measure-preserving (this follows from the ergodicity of the measure m_B^1 and the fact that π is non-singular).

Consider a factor map π: $\Sigma_A^h \to \Sigma_B^1$, where h has been adjusted so that π commutes with the flow. Now the flow on Σ_B^1 has eigenfrequency 1, and it follows that the flow on Σ_A^h also has eigenfrequency 1. By V.40 in [PT2] we may assume h is a positive rational-valued function which depends only on a finite number of coordinates, and for convenience we assume it depends only on the 0^{th} coordinate.

Put $h = k/n$, where n is a positive integer and k an integer-valued function. The Markov shift Σ_{A_k} may be viewed as the subspace $\{[x,i/n]: 0 \leq i < k(x_0)\}$ of Σ_A^h, equipped with the transformation $\sigma_A^{1/n}$; in fact $(\Sigma_A^h, \sigma_A^{1/n})$ may be identified (topologically and measure-theoretically) with $(\Sigma_{A_k} \times [0,1/n), \sigma_{A_k} \times id)$ (though we have to slightly adapt the topology of the latter space).

Similarly identify $(\Sigma_B^1, \sigma_B^{1/n})$ with $(\Sigma_{B_n} \times [0,1/n), \sigma_{B_n} \times id)$. Thus we obtain a m.p. semiconjugacy $(\Sigma_{A_k} \times [0,1/n), \sigma_{A_k} \times id) \xrightarrow{\pi} (\Sigma_{B_n} \times [0,1/n), \sigma_{B_n} \times id)$. By considering a dense orbit $\{\sigma_{A_k}^m(x)\}_{m \in \mathbb{Z}}$ we see that the image of a fibre $\Sigma_{A_k} \times \{s\}$ is contained in (and in fact is equal to) a fibre $\Sigma_{B_n} \times \{\alpha(s)\}$. By construction, α is a translation mod.$1/n$, and we may assume α is the identity.

For each $s \in [0,1/n)$ let π_s: $\Sigma_{A_k} \to \Sigma_{B_n}$ be defined by $\pi(x,s) = (\pi_s(x),s)$. Again because π is flow-commuting we conclude that $\pi_s = \pi_0$ for all s; hence $\pi = \pi_0 \times id$ and π_0: $\Sigma_{A_k} \to \Sigma_{B_n}$ is a measure preserving semiconjugacy. We have proved the following:

3.1. Proposition. ([PSu], [PT2]). *Let A and B be irreducible exponential matrices and π: $\Sigma_A^1 \to \Sigma_B^1$ a factor map. Then, for some integer-valued function k and some integer n, there exists a measure-preserving semiconjugacy π_0: $\Sigma_{A_k} \to \Sigma_{B_n}$, which is finite to one (resp. a.e. one to one, homeomorphism) if and only if π is finite to one (resp. a.e. one to one, homeomorphism).*

We remark that the SFT Σ_{A_k} is defined for every positive integer-valued continuous function k - although we only speak of the matrix A_k when k depends on a single coordinate.

3.2. Corollary [PT1]. *Stochastic flow equivalence (between suspensions of irreducible Markov shifts) is generated by expansion and block isomorphism.*

4. *The Δ and Γ groups.*

These groups are related to certain groups of non-singular self-transformations of Markov shifts and were first introduced in [K1] and [PSc]. Here we need only their definitions and a few basic results. The Δ and Γ groups are defined as follows:

$$\Delta_A = \{\frac{wt_A(u)}{wt_A(v)}: u, v \text{ cycles}, l(u) = l(v)\},$$

$$\Gamma_A = <\{wt_A(u): u \text{ cycle}\}>, \tag{3}$$

where <T> is the multiplicative group generated by the set T. Often we will drop the subscript A in our notation of these groups.

4.1. Proposition. [PSc]. *The group Γ/Δ is cyclic. Any irreducible exponential matrix is cometric with a matrix with entries in $\mathbb{Z}^+(\Delta^t)$ and conjugate to another with entries in $\mathbb{Z}^+(\Gamma^t)$.*

Note that the Δ-group of two cometric matrices is the same, whereas the Γ-group is preseved by conjugacy only.

4.2. Corollary. *A is a matrix of maximal type (i.e., m_A is a measure of maximal type) if and only if $\Delta = \{1\}$.*

Proof. If A is of maximal type then it is cometric with an integer matrix, which has

$\Delta = \{1\}$. It follows that $\Delta_A = \{1\}$.

Conversely, if Δ_A is the trivial group then A is cometric with a matrix which has entries in $\mathbb{Z}^+(1)$ - i.e. to an integer matrix. Hence A is of maximal type.□

The next proposition is a particularly simple case of a result in [PSc]. For the reader's convenience we include a self-contained proof.

4.3. Proposition. *If $\pi: \Sigma_A \to \Sigma_B$ is an a.e. one to one factor map then $\Delta_A = \Delta_B$, $\Gamma_A = \Gamma_B$.*

Proof. 2.4 implies that $\Delta_A \subseteq \Delta_B$, $\Gamma_A \subseteq \Gamma_B$. We now prove the other inclusions. By 2.1 there are a path $e_0 \ldots e_s$ in B, a number $0 \le j \le s$, and a vertex $J \in V_A$ such that, for every x $\in \pi^{-1}({}_0[e_0 \ \ e_s])$, $i(x_j) = J$.

Consider the set U of cycles $u = f_0 \ldots f_r$ in B such that $f_0 \ldots f_{s-j} = e_j \ldots e_s$ and $f_{r-j+1} \ldots f_r = e_0 \ldots e_{j-1}$: if we restrict the definitions (3) above to cycles u, v $\in$ U we still obtain the same Δ and Γ groups; and, since for each cycle u $\in$ U there is a cycle v in A with the same length as u and such that the periodic point u^∞ has image v^∞ (and therefore by 2.2 $wt_A(v) = wt_B(u)$), we obtain $\Delta_B \subseteq \Delta_A$ and $\Gamma_B \subseteq \Gamma_A$ as we wanted to.□

5. *Almost stochastic flow equivalence.*

The exponential matrices A and B are *almost stochastically flow equivalent* (a.s.f.e.) if there exists an exponential matrix C such that both $\Sigma_A^{|}$ and $\Sigma_B^{|}$ are a.e. one to one factors of $\Sigma_C^{|}$. The reader may check that the Γ-group is unchanged under expansion; combining this with propositions 4.3 and 3.1 we see that Γ is an invariant of almost stochastic flow equivalence.

We now come to our first main result.

5.1. Theorem. *The Γ-group is an invariant of a.s.f.e.; and any two irreducible exponential* *matrices with the same cyclic Γ-group are a.s.f.e.*

We shall need the following proposition.

5.2. Proposition. *Let A be an irreducible exponential matrix. Then A is flow equivalent to a matrix of maximal type if and only if its Γ-group is cyclic.*

We now complete the proof of the theorem, proving the proposition afterwards. Let A' and B' have the same cyclic Γ-group. By 5.2 we find matrices of maximal type A and B flow equivalent respectively to A' and B'. If necessary replacing A with A_l and B with B_n, for suitable positive integers l and n, we may assume that A and B have the same period d. If log β_A is the entropy of Σ_A then $\Gamma = <\beta_A^d>$ and similarly $\Gamma = <\beta_B^d>$ (this is where we use the fact that A and B are of maximal type). Therefore Σ_A and Σ_B have the same topological entropy and period, and by 2.2 they are almost conjugate. Since the maps involved preserve measures of maximal entropy, Σ_A and Σ_B are almost block isomorphic. By lifting the maps on the base spaces to the suspensions, this implies that A and B are a.s.f.e. - and so are A' and B'.

Proof. The "only if" is a consequence of 4.2 and 4.1. We prove the "if": by 4.2, a sufficient condition for an exponential matrix A to be of maximal type is that there be a $\gamma < 1$ such that, for every cycle u, the following equality holds:

$$wt_A(u) = \gamma^{l(u)} \tag{4}.$$

5.3. Lemma. *Let Σ_A be an irreducible SFT, and f: $\Sigma_A \to \mathbb{R}$ be a Holder-continuous function such that, for some positive constant γ, $\sum_{i=0}^{n-1} f(\sigma^i x) \geq \gamma n$ whenever $\sigma^n x = x$. Then f is cohomologous to a strictly positive continuous function.*

Proof. We assume familiarity with the theory of pressure and the Ruelle operator (see

[B2] or [R]). From the equality

$$P(tf) = \lim_{n \to +\infty} \frac{1}{n} \log \sum_{\sigma^n x = x} e^{tf^n(x)}$$

(where $t \in \mathbf{R}$, $P(tf)$ is the pressure of the function tf, and $f^n(x) = \sum_{i=0}^{n-1} f(\sigma^i x)$) we conclude, using our hypothesis, that $P(tf) \geq P(sf+(t-s)\gamma) = P(sf) + (t-s)\gamma$ whenever $t > s$. This shows at once the continuous function $t \to P(tf)$ is unbounded (both from above and from below) and strictly monotonic. Consider the unique $s < 0$ such that $P(sf) = 0$: by Ruelle's Perron-Frobenius theorem sf is cohomologous to a normalized function depending only on future coordinates - i.e. to a function g such that

$$\sum_{\{e:\, ex_0x_1\ldots \in \Sigma_A^+\}} e^{g(ex_0x_1\ldots)} = 1$$

for every $(x_0x_1\ldots) \in \Sigma_A^+$. This equality implies g is non-positive. Now f is cohomologous to g/s, which is a non-negative function. Since $f-\epsilon$ (for small $\epsilon > 0$) still satisfies the hypothesis of the lemma, it is also cohomologous to a non-negative function h_ϵ. Then f is cohomologous to $h_\epsilon+\epsilon$, which is strictly positive.□

We remark that examples of Holder-continuous functions on SFT are the locally constant functions - and, since every continuous function on an SFT can be aproximated by a locally constant function, 5.3 aplies to arbitrary continuous functions.

5.4. Lemma. *Let f be continuous and integer-valued (for convenience we assume it depends only on one coordinate). If $f(e_0) + \ldots + f(e_{n-1}) > 0$ for any cycle $e_0 \ldots e_{n-1}$ then f is cohomologous to a strictly positive, rational-valued, and locally constant function.*

Proof. We choose γ to be any positive number such that $f(e_0) + \ldots + f(e_{n-1}) \geq \gamma n$ for all *simple* cycles $e_0\ldots e_{n-1}$, so that γ satisfies the condition on 5.3. Then by 5.3 there is a continuous function k such that $g = f + k\circ\sigma - k$ is strictly positive. Replacing k with a sufficiently close rational-valued and locally constant function, the resulting g will be the function we want.□

Returning to the proof of 5.2, we assume A has cyclic Γ-group with generator $\beta > 1$. By 4.1 A is conjugate to a matrix $\tilde{A}$ such that $wt_{\tilde{A}}(e) = \beta^{-f(e)}$, where f is an integer-valued function on the edges of $\tilde{A}$. The function f satisfies the condition on 5.4, and so is cohomologous to a positive rational-valued function g which depends only on a finite number of coordinates. By going to higher blocks we then obtain a matrix $\tilde{A}$ of the form

$$\begin{bmatrix} \beta^{-g(1)} & & \\ & \ddots & \\ & & \beta^{-g(k)} \end{bmatrix}^{t} \mathbf{B},$$

where B is a 0-1 matrix. Put $g = h/n$, where n is integer and h is integer-valued. $\tilde{A}_h$ is a matrix of maximal type, since condition (4) holds with $\gamma = \beta^{-1/n}$.$\Box$

Remark. The above proof of 5.2 was suggested by Bill Parry. The author's original proof was rather long, and was based on the use of state-splitting.

6. *The topological case.*

The integer matrices A and B are *almost flow equivalent* if there exists an integer matrix C such that both Σ_A^1 and Σ_B^1 are a.e. one to one factors of Σ_C^1. We now state and prove Boyle's theorem, which turns out to be a simple consequence of our theorem 5.1.

6.1. Theorem (Boyle [B3]). *Any two irreducible SFT with positive topological entropy are almost flow equivalent.*

We have shown (5.1) that two irreducible Markov shifts with the same cyclic Γ-group are almost stochastically flow equivalent, and so our proof is based on the construction of Markov measures with specified Γ-group on arbitrary SFT.

6.2. Lemma. *For every irreducible SFT Σ_A with positive topological entropy there exists a fully supported Markov measure μ on Σ_A such that $\Gamma(\mu) = <1/2^k>$ (for some $k \geq 1$).*

Proof. We define a weight function wt as follows: let $I \in V_A$; if there are k_I edges going out of I, we number these edges as $e_{I,i}$, $1 \leq i \leq k_I$, and define $wt(e_{I,i}) = 1/2^i$ $(1 \leq i < k_I)$ and $wt(e_{I,k_I}) = 1/2^{k_I - 1}$. The matrix $A_{wt}(1)$ is stochastic, and so A_{wt} defines a Markov measure μ with $\Gamma(\mu) = <1/2^k>$, some k, for all transition probabilities are powers of 1/2.□

6.3. Lemma. *For every $k \geq 1$ there exists an irreducible SFT Σ_k having fully supported Markov measures μ_k, m_k such that $\Gamma(\mu_k) = <1/2^k>$ and $\Gamma(m_k) = <1/2>$.*

Proof. Take Σ_k to be the full shift on 2^k symbols, μ_k and m_k to be the Bernoulli shifts based respectively on the probabiblity vectors $(1/2^k, 1/2^k, \ldots, 1/2^k)$ and $(1/2, 1/4, \ldots, 1/2^{n-1}, 1/2^{n-1})$, where $n = 2^k$.□

Proof of theorem 6.1. Let Σ_A be an irreducible SFT with positive topological entropy. Using 6.2, endow it with a Markov measure μ such that $\Gamma(\mu) = <1/2^k>$: by 5.1, (Σ_A, μ) and (Σ_k, μ_k) are a.a.f.e. - hence Σ_A and Σ_k are almost flow equivalent.

For any two integers k, k' $\geq$ 1, (Σ_k, m_k) and $(\Sigma_{k'}, m_{k'})$ are a.a.f.e. (since $\Gamma(m_k) = \Gamma(m_{k'}) = <1/2>$), and so Σ_k and $\Sigma_{k'}$ are almost flow equivalent - which implies that the same holds for any two irreducible SFT with positive entropy.□

Remark. In [B3] Boyle proves a more general result than 6.1: if A and B are irreducible integer matrices then the suspension Σ_A^1 is a finite to one factor of Σ_B^1 iff the Bowen-Franks' group BF(A) is a homomorphic image of BF(B); the factor map $\Sigma_A^1 \to \Sigma_B^1$ can be constructed as a.e. one to one and - if det(I-B) = 0 or the signs of det(I-A) and det(I-B) are equal - as a bijection on all but finitely many periodic orbits.

7. *Hyperbolic flows.*

Here we briefly describe hyperbolic flows, and how to translate theorem 6.1 to this setting. A helpful reference for this section is [B1].

Let M be a compact Riemann manifold and φ_t: M → M a differentiable flow. A closed invariant set Λ is *hyperbolic* if it contains no fixed points and there is a continuous decomposition of the tangent bundle restricted to Λ into a Whitney sum $T_\Lambda M = E \oplus E^u \oplus E^s$ of T_φ-invariant sub-bundles with the following properties:

(i) E_x is the one-dimensional space tangent to the flow, for each $x \in \Lambda$;

(ii) there are positive constants c, λ such that $\| T\varphi_t(v) \| \leq ce^{-\lambda t} \| v \|$ for all $t \geq 0$ and $v \in E^s$, and $\| T\varphi_{-t}(v) \| \leq ce^{-\lambda t} \| v \|$ for all $t \geq 0$ and $v \in E^u$.

A hyperbolic set $\Lambda \subset M$ is called a *hyperbolic basic set* if it has the following properties:

(i) $\varphi_t|_\Lambda$ is transitive, i.e. it possesses a dense orbit;

(ii) the periodic orbits of $\varphi_t|_\Lambda$ form a dense subset of Λ;

(iii) there exists an open set $U \supset \Lambda$ such that $\Lambda = \bigcap_{t \in \mathbb{R}} \varphi_t(U)$.

An *Axiom A* flow is a flow whose non-wandering set is both a hyperbolic set and the closure of the periodic orbits. Then the non-wandering set is a finite union of (necessarily disjoint) hyperbolic basic sets. An Axiom A flow whose non-wandering set is one-dimensional is called a *Smale flow.*

The following theorem establishes the connection between hyperbolic basic sets and symbolic flows:

7.1. Theorem (Bowen [B1]). *If* Λ *is a hyperbolic basic set of the flow* φ_t *then there exists a* *symbolic flow* Σ^f_A *and a function* π: $\Sigma^f_A \to \Lambda$ *with the properties:*

(i) π *is continuous;*

(ii) $\pi \circ \sigma^t_A = \varphi_t \circ \pi$ *for all* $t \in \mathbf{R}$;

(iii) π *is surjective and bounded to one;*

(iv) π *is one-one on a set of full measure (for any fully supported ergodic measure) and on a Baire set of first category.*

If furthermore Λ *is one-dimensional then* π *can be chosen to be a bijection - i.e., a conjugacy.*

Theorem 6.1 has then the following easy consequence, which was actually a strong motivation for the original problem:

7.2. Theorem (Boyle [B3]). *Let* Λ_1 *and* Λ_2 *be hyperbolic basic sets of the flows* φ_t *and* ψ_t, *and assume* Λ_1 *and* Λ_2 *do not consist of a single closed orbit. Then there exists a symbolic flow* Σ^1_A *such that there are functions* π_1: $\Sigma^1_A \to \Lambda_1$ *and* π_2: $\Sigma^1_A \to \Lambda_2$ *both with properties (i), (iii), (iv) above and, instead of (ii), the property:*

(ii') π_1 and π_2 send orbits onto orbits, preserving the orientation, and their restrictions to each orbit are local homeomorphisms onto its images.

A *continuous reparametrization* of the flow φ_s on Λ is a flow $\psi_t(x) = \varphi_{s(x,t)}(x)$, where each function $s(x,.)$: $\mathbf{R} \to \mathbf{R}$ is an orientation-preserving homeomorphism, and the function $(x,t) \mapsto s(x,t)$ is continuous and satisfies the equality $s(x,t+r) = s(x,r) + s(\varphi_{s(x,r)}(x), t)$ for every $t, r \in \mathbf{R}$.

The next theorem specialises and refines the preceding one for the case of one-dimensional flows. It follows from the proof of 6.1, as an examination of it easily shows.

7.3. Theorem. *Let* φ_s *and* ψ_s *be flows on one-dimensional hyperbolic basic sets (e.g., basic sets of Smale flows)* Λ_1 *and* Λ_2 *(not consisting of a single closed orbit). Then there exist* *continuous reparametrizations φ_t of $\varphi_s|_{\Lambda_1}$ and ψ_t of $\psi_s|_{\Lambda_2}$ such that there are a symbolic flow Σ_A^f and functions π_1: $\Sigma_A^f \to \Lambda_1$ and π_2: $\Sigma_A^f \to \Lambda_2$ satisfying properties (i) - (iv) of 7.1 with respect to the flows φ_t and ψ_t.*

8. *Examples.*

We present here two examples, the first illustrating a method of constructing on a given irreducible SFT a Markov measure with Γ-group generated by $1/\beta$, for certain Perron numbers β, the other showing that the second statement of theorem 5.1 can not be generalized to arbitrary Γ-groups.

We begin with a lemma which proves to be very useful to both examples.

8.1. Lemma. *Let A be an irreducible integer matrix, and let Q be an exponential matrix compatible with A (i.e., $Q(0) = A$). Given a state $I \in V_A$ let $\mathcal{A}(I)$ be the set of loops $c = e_0 \ldots e_n$ that begin and end at I and such that $i(e_j) \neq I$ for $1 \leq j \leq n$. The following statements are equivalent:*

(i) *$Q(1)$ has spectral radius 1;*

(ii) *there exists $I \in V_A$ such that $\sum_{c\in\mathcal{A}(I)} wt_Q(c) = 1$;*

(iii) *$\sum_{c\in\mathcal{A}(I)} wt_Q(c) = 1$ for all $I \in V_A$.*

Proof. Let $P = \lambda^t D^t Q D^{-t}$ be the unique matrix cometric with Q such that P(1) is stochastic, m_P the corresponding Markov measure, and p the probability vector such that pP(1) = p. Modulo a set of m_P-measure zero, the set $_0[I]$ can be partitioned into a disjoint union $\bigcup_{c\in\mathcal{A}(I)} {}_0[c]$ - i.e., $m_P({}_0[I] \setminus \bigcup_{c\in\mathcal{A}(I)} {}_0[c]) = 0$. We therefore have

$$p(I) = m_P({}_0[I]) = \sum_{c\in\mathcal{A}(I)} m_P({}_0[c]) = p(I) \sum_{c\in\mathcal{A}(I)} wt_P(c),$$

which implies that

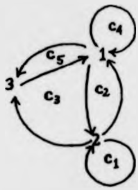
$$\sum_{c \in \mathcal{A}(l)} wt_P(c) = 1.$$

To prove the lemma we show (i) $\Rightarrow$ (iii) $\Rightarrow$ (ii) $\Rightarrow$ (i). Assuming (i) holds, we have $\lambda = 1$ and therefore $\sum_{c \in \mathcal{A}(l)} wt_Q(c) = \sum_{c \in \mathcal{A}(l)} wt_P(c) = 1$ for any l, which is condition (iii). The implication (iii) $\Rightarrow$ (ii) is obvious. If we assume (ii) holds then for some vertex l we have $\sum_{c \in \mathcal{A}(l)} wt_Q(c) = \sum_{c \in \mathcal{A}(l)} wt_P(c)$, an equality which is possible only if $\lambda = 1$, and this means Q(1) has spectral radius 1. We have proved (ii) $\Rightarrow$ (i).□

Example 1. Consider the matrix

$$A = \begin{bmatrix} 1 & 1 & 1 \\ 1 & 1 & 1 \\ 1 & 0 & 0 \end{bmatrix}$$

and the associated graph represented below, where the symbols $c_1, \ldots, c_5$ designate the simple cycles:



The characteristic polynomial of A is $p_A(x) = x\,(x^2 - 2x - 1)$. Our aim is to construct a fully supported Markov measure μ on Σ_A with $\Gamma(\mu) = <1/\beta>$, where β is the largest root (in absolute value) of the polynomial $q(x) = x^3 - 3x^2 + 2x - 1$. This is a non-trivial problem.

We treat $c_1, \ldots, c_5$ as positive real variables, and their possible values as weights of the corresponding cycles. The equation $\sum_{c \in \mathcal{A}_i} wt_Q(c) = 1$ in 8.1 can then be rewritten as

$$c_4 + (c_2 + c_3)(1 + c_1 + c_1^2 + \dots) + c_5 = 1$$

or

$$1 - (c_1 + c_2 + c_3 + c_4 + c_5) + c_1 c_4 + c_1 c_5 = 0 \qquad (5).$$

We call equation (5) the *loop equation*, the polynomial $r_A(c_1, \dots, c_5)$ on the left-hand side the *loop polynomial*, and the integer $n_A = r_A(1, \dots, 1)$ the *loop number*.

If γ is the maximum eigenvalue of A and we put $c_i = \gamma^{-l(c_i)}$ then the resulting quintuple $(c_1, \dots, c_5)$ is a solution of the loop equation. Since $-x^3 p_A(x^{-1})$ is the minimum polynomial of γ^{-1}, it has to divide the polynomial $r_A(x, x^2, x^3, x, x^2)$.

Let $p(x) = -x^3 q(x^{-1})$ be the minimum polynomial of $1/\beta$. What we are then trying to find are integer numbers $n_i \geq 1$ such that $p(x)$ divides $r(x^{n_1}, \dots, x^{n_5})$. One necessary condition for the existence of such numbers is that $p(1)$ divides n_A.

This necessary condition is met, because $p(1) = -q(1) = 1$. We multiply $p(x)$ by a polynomial $s(x)$ with $s(1) = -2$ so that $(sp)(1) = n_A$. Our choice is $s(x) = -(x + 1)$, and we obtain $(sp)(x) = -x^4 + x^3 - x^2 - 2x + 1$. We then note that $(sp)(x) = r_A(x, x^4, x^2, x, x^2)$.

Thus the Markov measure μ we are searching for is constructed by attributing the weights β^{-1}, β^{-4}, β^{-2}, β^{-1}, β^{-2} respectively to the cycles c_1, c_2, c_3, c_4, c_5. One possibility is to define μ by the matrix

$$Q = \begin{bmatrix} \beta^{-1} & 1 & \beta^{-2} \\ \beta^{-4} & \beta^{-1} & \beta^{-2} \\ 1 & 0 & 0 \end{bmatrix}.$$

In the next section we shall give a proper definition of the loop polynomial, and shall introduce a combinatorial machinery that proves adequate to analize the loop structure of irreducible directed graphs. The immediate motivation for this analysis is of course to prove that the sums of the type $\sum_{c \in \mathcal{A}(l)} wt_P(c)$ have a simple expression in terms of the weights of the simple cycles.

The method we illustrated above can in principle be used to construct Markov measures with more general (i.e. not just cyclic) Γ-groups, albeit with complicated calculations. Theorem

9.6 tells us precisely the conditions on $wt_Q(c_1), \ldots, wt_Q(c_k)$ (where $c_1, \ldots, c_k$ are the simple cycles) under which we can ensure Q has spectral radius 1.

As the definition of loop polynomial will make clear, the loop number n_A is equal to the Parry-Sullivan invariant det(I-A). We conclude these remarks with a simple proposition:

8.2. Proposition. *One necessary condition for the existence of a Markov measure μ supported on Σ_A with $\Gamma(\mu) = <1/\beta>$, where $\beta > 1$, is that the loop number n_A be a multiple of $q(1)$, where $q(x)$ is the minimum polynomial of β.*

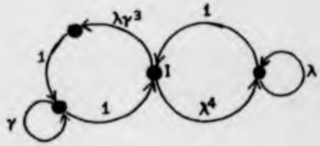
The proof rests on the observation that if $r_A(c_1, \ldots, c_k)$ is the loop polynomial of A then, if such measure μ exists, there are integers $n_i \geq 1$ such that $r_A(\beta^{-n_1}, \ldots, \beta^{-n_k}) = 0$, and therefore the minimum polynomial p of $1/\beta$ divides $r_A(x^{n_1}, \ldots, x^{n_k})$. It follows that p(1) divides n_A, and the proof is finished by noting that $p(1) = \pm q(1)$.

Example 2. Now we have an example of two irreducible Markov shifts which have the same Γ-group and cannot be almost stochastically flow equivalent, thereby disproving the natural conjecture that this generalization of the statement of 5.1 is true. First we present our example, and then we develop some theory, motivated by [MT], which allows us to show why our example works.

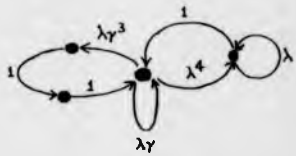
Let $\lambda, \gamma \in [0, 1]$ be defined by the equalities $\lambda^2 + \lambda = 1$ and $\gamma^3 + \gamma = 1$. The numbers λ^k and γ^l can only be in Q if they are both equal to 1, and therefore $Q(\lambda^k)$ and $Q(\gamma^l)$, for $k, l \neq 0$, are algebraic extensions of Q of degrees 2 and 3, respectively. This shows that the equality $\lambda^k = \gamma^l$, for $k, l \in \mathbb{Z}$, is only possible for $k = l = 0$.

We define the Markov shifts Σ_{P_1} and Σ_{P_2} by their graphs as follows.

P_1:



P_2:



Using 8.1 we may verify that both P_1 and P_2 have spectral radius 1. For instance, we calculate

$$\sum_{c\in\mathcal{A}(1)} wt_{P_1}(c) = \lambda\gamma^3(1+\gamma+\gamma^2+\dots)+\lambda^4(1+\lambda+\lambda^2+\dots)$$
$$= \lambda\frac{\gamma^3}{1-\gamma}+\lambda^2\frac{\lambda^2}{1-\lambda} = \lambda+\lambda^2 = 1,$$

and similarly for P_2. Also note that $\Gamma_{P_1} = \Gamma_{P_2} = \langle\lambda, \gamma\rangle$.

We now introduce a couple of definitions. Let P be an irreducible exponential matrix such that P(1) has spectral radius 1. If c is a cycle in P, the *weight-per-symbol* of c is $wps(c) = \frac{1}{l(c)} \log wt_P(c)$. The *weight-per-symbol set* WPS(P) is the collection of all wps(c) of cycles c in P.

8.3. Lemma (Marcus and Tuncel [MT]). *If π: $\Sigma_P \to \Sigma_Q$ is a finite to one factor map then WPS(P) = WPS(Q).*

Proof. We reproduce here the very short proof for the reader's convenience. If c is a cycle in P then by 2.4 we have $wps(\pi(c)) = wps(c)$. Hence $WPS(P) \subseteq WPS(Q)$. To prove the other inclusion, consider a cycle d in Q. Then for some $k \geq 1$ there exists a cycle c in P such that $\pi(c) = d^k$, and it follows that $wps(c) = wps(d^k) = wps(d)$.□

If S a set of real numbers, let $<S>_{\mathbf{Q}^+}$ denote the set of linear combinations with non-negative rational coefficients of elements in S, and Ch(S) the convex hull of S over Q.

8.4. Lemma. *If P and Q are flow equivalent then* $<WPS(P)>_{\mathbf{Q}^+} = <WPS(Q)>_{\mathbf{Q}^+}$.

Proof. Note that WPS(P) $\subseteq$ Ch({wps(c): c is a simple cycle in P}) and therefore $<WPS(P)>_{\mathbf{Q}^+} = <\{wps(c): c \text{ simple cycle}\}>_{\mathbf{Q}^+}$. If we alter the graph P by expansions, thus obtaining a graph P, then there is a simple relation between the wps(c) of a simple cycle in P and the wps(c) of the corresponding cycle in P, namely that $wps(c) = \frac{l(c)}{l(\bar{c})}\, wps(c)$. It follows that $<WPS(P)>_{\mathbf{Q}^+} = <WPS(P)>_{\mathbf{Q}^+}$. Therefore the set $<WPS(P)>_{\mathbf{Q}^+}$ is not affected by expansions, nor is it affected by passing to a conjugate graph. The result now follows from 3.2.□

Combining 8.3 and 8.4 we obtain, with the help of proposition 3.1, the following important result:

8.5. Proposition. *If P and Q are almost stochastically flow equivalent then* $<WPS(P)>_{\mathbf{Q}^+} = <WPS(Q)>_{\mathbf{Q}^+}$.

Coming back to our example, we note that WPS(P_1), and therefore $<WPS(P_1)>_{\mathbf{Q}^+}$, is not contained in $<WPS(P_2)>_{\mathbf{Q}^+}$, for instance because the equality

$$\log \gamma = q_1(\tfrac{1}{3} \log \lambda\gamma^3) + q_2 \log \lambda\gamma + q_3 (\tfrac{1}{2} \log \lambda^4) + q_4 \log \lambda$$

is impossible for non-negative rational numbers $q_1, \ldots, q_4$. Therefore P_1 and P_2 are not almost stochastically flow equivalent.

9. *The loop polynomial.*

We met the loop polynomial in example 1 of the previous section. It results, after some manipulation, from the equation $\sum_{c \in \mathcal{A}(1)} wt_Q(c) = 1$ in lemma 8.1 However, it is not clear in general how to calculate this polynomial, or indeed whether it does always exist. Here we give its definition, which is related to a method of calculating det(I-A), for a given square matrix A, and prove a number of results on the loop structure of directed graphs which put the loop polynomial to good use.

If A is an irreducible integer matrix, and Q an exponential matrix, we say Q has loop structure A if $Q(0) = A$. Throughout this section, it is useful to think of A as the equivalence class of exponential matrices with loop structure A. For simplicity of notation we assume A is a 0-1 matrix, although our results are stated without this restriction.

A *simple cycle* (we shall for the moment abreviate this as *cycle*) in A is represented by a k-tuple $(I_1, \dots, I_k)$, where $1 \leq k \leq n$, $1 \leq I_s \leq n$ for each s, all I_s are distinct, and the entries $A(I_s, I_{s+1})$ $(1 \leq s < k)$, $A(I_k, I_1)$ are all non-zero. Two k-tuples $(I_1, \dots, I_k)$ and $(J_1, \dots, J_k)$ represent the same cycle if, for some $0 \leq t < k$ and all s, we have $J_s = I_{s+t \pmod k}$. Two cycles $c = (I_1, \dots, I_k)$ and $d = (J_1, \dots, J_l)$ are said to be *disjoint* if the sets $V_c = \{I_1, \dots, I_k\}$ and $V_d = \{J_1, \dots, J_l\}$ are disjoint; otherwise c and d are *neighbours*. The *length* of c, which we denote by $l(c)$, is the cardinality of the set V_c.

Each cycle $c = (I_1, \dots, I_k)$ defines a permutation $\sigma_c \in S_n$ which is the identity outside V_c and such that $\sigma_c(I_s) = I_{s+1 \pmod k}$. It is well known that $\mathrm{sgn}(\sigma_c) = (-1)^{l(c)-1}$ and that each permutation $\sigma \in S_n$ with $A(1, \sigma(1)) \dots A(n, \sigma(n)) \neq 0$ can be written uniquely (except for the order of the factors) as a composition $\sigma = \sigma_{c_1} \circ \dots \circ \sigma_{c_l}$, where the c_i are pairwise disjoint cycles and $l(c_1) + \dots + l(c_l) = n$.

Let us now enumerate the different cycles in A as $c_1, \dots, c_k$. If $1 \leq i_1 < \dots < i_s \leq k$, we define the symbol $\delta_{i_1 \dots i_s}$ as follows:

(i) $\delta_{i_1 \dots i_s} = 1$ if the cycles $c_{i_1}, \dots, c_{i_s}$ are pairwise disjoint and $l(c_{i_1}) + \dots + l(c_{i_s}) = n$;

(ii) $\delta_{i_1 \dots i_s} = 0$ otherwise.

We shall also make use of the symbol $\epsilon_{i_1 \ldots i_s}$, which is defined by:

(i) $\epsilon_{i_1 \ldots i_s} = (-1)^s$ if either $s = 0$ (i.e., we define $\epsilon_\emptyset = 1$, where the subscript is empty) or the cycles $c_{i_1}, \ldots, c_{i_s}$ are pairwise disjoint;

(ii) $\epsilon_{i_1 \ldots i_s} = 0$ otherwise.

Keeping in mind the preceding observations and the formula $\det A = \sum_{\sigma \in S_n} \operatorname{sgn}(\sigma)\, A(1, \sigma(1)) \ldots A(n, \sigma(n))$ we can write the following formulae:

$$\det A = \sum_{1 \le i_1 < \ldots < i_s \le k} \epsilon_{i_1 \ldots i_s} (-1)^{l(c_{i_1})-1} c_{i_1} \ldots (-1)^{l(c_{i_s})-1} c_{i_s} \tag{6}$$

$$\det (I - A) = \sum_{1 \le i_1 < \ldots < i_s \le k} \epsilon_{i_1 \ldots i_s} c_{i_1} \ldots c_{i_s} \tag{7}$$

where each $c_i = (J_1, \ldots, J_l)$ stands for the product $A(J_1, J_2) \ldots A(J_l, J_1)$.

The polynomial on the left-hand side of formula (7) is called the *loop polynomial* of A and will be denoted by $r_A(c_1, \ldots, c_k)$ (or simply r_A). The reader may check that this definition agrees with the polynomial we have computed in example 8.1.

If Q has loop structure A, it is obvious that the equality $r_A(wt_Q(c_1), \ldots, wt_Q(c_k)) = 0$ simply says Q has the eigenvalue 1. In order that a solution of the loop equation $r_A = 0$ correspond to a matrix of spectral radius 1, extra conditions have to be verified.

We now state the main result to be proved in this section.

9.1. Theorem. *For an exponential matrix Q with loop structure A, the following conditions are equivalent:*

(i) *Q satisfies the equality* $\sum_{c \in \mathcal{A}(I)} wt_Q(c) = 1$ *for some vertex I;*

(ii) $\sum_{c \in \mathcal{A}(I)} wt_Q(c)$ *is finite for all vertices I and the k-tuple* $(wt_Q(c_1), \ldots, wt_Q(c_k))$ *is a root of the loop polynomial* r_A.

At the end of the proof of 9.1 we shall have occasion to remark that the condition of $\sum_{c \in \mathcal{A}(I)} \mathrm{wt}_Q(c)$ being finite for all I can be expressed by finitely many polynomial inequalities. This provides a more manageable form of theorem 9.1, which we state as theorem 9.6.

Proof of 9.1. The proof consists of writing $\sum_{c \in \mathcal{A}(I)} \mathrm{wt}_Q(c)$ as a (formal) power series on the variables $c_1, \ldots, c_k$ and working out its sum. Only at the end shall we worry about the convergence of this and other series.

First we explain how to decompose a given cycle c based at I into a "product" of simple cycles. We put $c = I_0 I_1 \ldots I_r$, where $I_0 = I_r = I$. If c is not a simple cycle we let $t \geq 1$ be the maximum index such that I_t occurs in $I_1 \ldots I_r$ at least twice. We can then write $c = I\, U\, I_t\, V\, I_t\, W$, where U, V, W are (perhaps empty) sequences of symbols, and I_t occurs in neither V nor W. Furthermore, there are in c no repetitions of symbols in W. If we let $u = I\, U\, I_t\, W$, $v = I_t\, V\, I_t$, which are cycles based respectively at I and I_t, and assume as an inductive step we have already decomposed u as a product $c_{i_0}\, c_{i_1} \ldots c_{i_l}$, and v as a product $c_{j_0}\, c_{j_1} \ldots c_{j_m}$, then the decomposition of c is the product $c_{i_0}\, c_{i_1} \ldots c_{i_l}\, c_{j_0}\, c_{j_1} \ldots c_{j_m}$. In this decomposition, or *factorization*, the order of the factors is crucial.

If c is not simple, the cycle v that we cut off from c to obtain u is called a *deleting loop* of c; the deleting loops of c comprise, in addition to v, all the deleting loops of u. Thus, if $v_1, \ldots, v_r$ is the enumeration of the deleting loops of c by their order of occurrence (i.e., $v_r = v$), and if after cutting off all the v_i from c we obtain c_{i_0}, then the factorization of c is the concatenation of c_{i_0} with the factorizations of $v_1, \ldots, v_r$ (in this order).

9.2. *Proposition. There is a one to one correspondence between the cycles based at I and their factorizations.*

Proof. We consider the factorisation $c_{i_0}\, c_{i_1} \ldots c_{i_l}\, c_{j_0}\, c_{j_1} \ldots c_{j_m}$ as given above, and write $c_{i_0} = J_0 \ldots J_s$, where $J_0 = J_s = I$. If r is the maximum index such that J_r also occurs in some other cycle c_i $(i = i_1, \dots, i_l, j_0, \dots, j_m)$, then c_{j_0} is the last such cycle where J_r occurs. This selection of J_r and c_{j_0} depends only on the factorization itself. We then assume, by induction, that $c_{i_0} c_{i_1} \dots c_{i_l}$ and $c_{j_0} c_{j_1} \dots c_{j_m}$ are the factorizations of unique cycles u based at I and v based at J_r. If we write $v = J_r\ V\ J_r$, $u = I\ U\ J_r\ W$ (where J_r does not occur in W), then $c = I\ U\ J_r\ V\ J_r\ W$ is the unique cycle whose factorization is $c_{i_0} c_{i_1} \dots c_{i_l} c_{j_0} c_{j_1} \dots c_{j_m}$.□

The main point is that it now makes sense to say that the simple cycle c_i occurs in c (it does so if it occurs in the factorization of c), and to count the number of times it occurs. And to obtain the formula for the sum $\sum_{c \in \mathcal{A}(I)} wt_Q(c)$ we replace each $wt_Q(c)$ with $c_1^{m_1(c)} \dots c_k^{m_k(c)}$, where $m_i(c)$ is the number of times c_i occurs in c.

Examples. Consider the directed graphs A and B in the figure below. We have the following examples of factorizations of cycles in A:

$1\ 2\ 2\ 3\ 3\ 2\ 3\ 1 \sim c_1 c_2 c_4 c_3$,

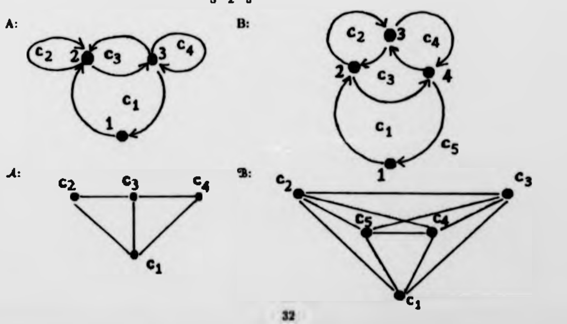
$1\ 2\ 2\ 3\ 2\ 3\ 3\ 1 \sim c_1 c_2 c_3 c_4$,

$1\ 2\ 3\ 2\ 2\ 3\ 3\ 1 \sim c_1 c_3 c_2 c_4$;

and the following examples of cycles in B:

$1\ 2\ 3\ 2\ 4\ 3\ 2\ 4\ 1 \sim c_1 c_2 c_3$,

$1\ 2\ 3\ 2\ 3\ 4\ 3\ 2\ 4\ 1 \sim c_5 c_2 c_3$.



We introduce now the loop diagram, which contains, as we shall see, all the information we need on the loop structure of a given directed graph. We have represented in the above figure the loop diagrams $\mathcal{A}$ and $\mathfrak{B}$ corresponding to the graphs A and B.

The *loop diagram* of a given irreducible directed graph A is an undirected graph $\mathcal{A}$ where the nodes are the simple cycles in A, and there is an arc joining two nodes if the cycles they represent are neighbours. In general any subdiagram $\mathfrak{B}$ of a loop diagram $\mathcal{A}$ (which is defined by taking for nodes a subset of the nodes of $\mathcal{A}$, and for arcs the arcs that join these nodes in $\mathcal{A}$) is called a loop diagram. $\mathcal{A}\backslash S$ indicates the subdiagram of $\mathcal{A}$ which contains all the nodes in $\mathcal{A}$ except those in the set S. Note that a loop diagram is not necessarily the full loop diagram of any directed graph: if $\mathcal{A}$ is the loop diagram of the graph A, and $\mathfrak{B}$ is a subdiagram of $\mathcal{A}$, then it does not follow $\mathfrak{B}$ is the loop diagram of some subgraph of A.

Let c be a simple cycle and I a vertex in c. Define $\mathcal{A}_c(I)$ as the set of loops based at I whose factorizations begin with c and which don't have any further occurrences of c.

If $\mathcal{A}$ is the loop diagram of A, $\mathfrak{B}$ is a subdiagram of $\mathcal{A}$, and c is a node in $\mathfrak{B}$, we define

$$(c)_{\mathfrak{B},I} = \sum_{d \in \mathfrak{B}_c(I)} wt_Q(d),$$

where $\mathfrak{B}_c(I)$ is the set of those cycles in $\mathcal{A}_c(I)$ whose factorizations contain only cycles in $\mathfrak{B}$. If $\mathfrak{B}$ is not connected then we have $(c)_{\mathfrak{B},I} = (c)_{\mathfrak{C},I}$, where $\mathfrak{C}$ is the connected component of $\mathfrak{B}$ containing c.

The proof of the following crucial lemma is given at the end of this section.

9.3. Lemma. *If both I and J are vertices in c then $(c)_{\mathfrak{B},I} = (c)_{\mathfrak{B},J}$.*

Henceforth we use the notation $(c)_{\mathfrak{B}}$ instead of $(c)_{\mathfrak{B},I}$.

In our first example, let $\mathfrak{C}$ be the subdiagram of $\mathcal{A}$ containing the nodes c_1, c_2, c_3, and $\mathfrak{B}$ the one containing the nodes c_2, c_3, c_4. If d_1 is a loop in $\mathfrak{C}_{c_1}(1)$ and $d_2, d_3, \ldots, d_n$ are loops in $\mathfrak{B}_{c_4}(3)$ then the cycle whose factorization is the concatenation of the factorizations of $d_1, \ldots, d_n$ (in this order) belongs to $\mathcal{A}_{c_1}(1)$. It is not difficult to see that all the cycles in $\mathcal{A}_{c_1}(1)$ are obtained in this way. We therefore have

$$(c_1)_{\mathcal{A}} = (c_1)_{\mathcal{C}} [1 + (c_4)_{\mathfrak{B}} + (c_4)_{\mathfrak{B}}{}^2 + (c_4)_{\mathfrak{B}}{}^3 + \dots]$$

$$= \frac{(c_1)_{\mathcal{C}}}{1 - (c_4)_{\mathfrak{B}}} \qquad (8).$$

Formula (8) is very important, providing the induction step in the proof of our main lemma 9.5. Before turning to the general case, we conclude our calculation by observing that $(c_1)_{\mathcal{C}} = \frac{c_1}{1 - (c_2 + c_3)}$, $(c_4)_{\mathfrak{B}} = \frac{c_4 - c_2 c_4}{1 - (c_2 + c_3)}$, and (8) then gives

$$(c_1)_{\mathcal{A}} = \frac{c_1}{1 - (c_2 + c_3 + c_4) + c_2 c_4}.$$

9.4. Lemma. *If $\mathfrak{D}$ is a loop diagram and c is a node in $\mathfrak{D}$ with some neighbour in $\mathfrak{D}$, then there exists in $\mathfrak{D}$ a neighbour d of c such that* $(c)_{\mathfrak{D}} = \dfrac{(c)_{\mathfrak{D}\setminus\{d\}}}{1 - (d)_{\mathfrak{D}\setminus\{c\}}}$.

Proof. Fix a vertex I in c and write $c = I_0 \dots I_r$, where $I_0 = I_r = I$. Let $t \geq 1$ be the maximum index such that I_t belongs to some other cycle in $\mathfrak{D}$, and choose $d \in \mathfrak{D}\setminus\{c\}$ as one of the cycles which contain I_t. If $d_1 = I\ V_1\ I_t\ W$ (where W may be empty and I_t does not occur in W) is a cycle in $(\mathfrak{D}\setminus\{d\})_c(I)$, and $d_i = I_t\ V_i\ I_t$ $(2 \leq i \leq s,\ s \geq 1)$ are cycles in $(\mathfrak{D}\setminus\{c\})_d(I_t)$, then $I\ V_1\ I_t\ V_2\ I_t \dots I_t\ V_s\ I_t\ W$ is the cycle in $\mathfrak{D}_c(I)$ whose factorization is the concatenation of the factorizations of $d_1, \dots, d_s$ (in this order). Since each cycle in $\mathfrak{D}_c(I)$ is obtained exactly once in this way, we repeat the calculation in (8) to arrive at the desired conclusion.□

Let $\mathfrak{D}$ be a loop diagram . A *separation of* $\mathfrak{D}$ is a partition of $\mathfrak{D}$ with only two sets $\mathfrak{C}$ and $\mathfrak{D}$, and such that, for $c \in \mathfrak{C}$ and $d \in \mathfrak{D}$, c and d are never neighbours. Let $c_1, \dots, c_t$ be an enumeration of the nodes in $\mathfrak{D}$. We define

$$[\mathfrak{D}] = \sum_{1 \leq i_1 < \dots < i_s \leq t} \epsilon_{i_1 \dots i_s}\, c_{i_1} \cdots c_{i_s}$$

(for the empty diagram we define $[\emptyset] = 1$), and denote by $\mathbb{Z}(\mathfrak{B})$ the field of fractions of integer polynomials on the variables $c_1, \ldots, c_\ell$. Note that $[\mathcal{A}] = r_A$ if $\mathcal{A}$ is the loop diagram of A. We also remark that $[\mathfrak{B}] = [\mathfrak{C}]\,[\mathfrak{D}]$ when $\mathfrak{C}$ and $\mathfrak{D}$ form a separation of $\mathfrak{B}$.

9.5. Lemma. *Let $\mathcal{A}$ be a loop diagram, and suppose we assign to each subdiagram $\mathfrak{B}$ of $\mathcal{A}$, and each node c in $\mathfrak{B}$, an element $(c)_{\mathfrak{B}}$ in $\mathbb{Z}(\mathfrak{B})$. Assume this assignment has the following properties:*

(i) $(c)_{\{c\}} = c$;

(ii) *if c is a node in $\mathfrak{B}$ and $\mathfrak{C}$ is the connected component of $\mathfrak{B}$ which contains c then* $(c)_{\mathfrak{B}} = (c)_{\mathfrak{C}}$;

(iii) *for all c in $\mathfrak{B}$ there exists in $\mathfrak{B}$ a neighbour d of c such that*

$$(c)_{\mathfrak{B}} = \frac{(c)_{\mathfrak{B}\setminus\{d\}}}{1 - (d)_{\mathfrak{B}\setminus\{c\}}} \tag{9}$$

Then we have

$$(c)_{\mathfrak{B}} = \frac{[\mathfrak{B}\setminus\{c\}] - [\mathfrak{B}]}{[\mathfrak{B}\setminus\{c\}]} \tag{10}$$

for all subdiagrams $\mathfrak{B}$ and all c in $\mathfrak{B}$. Furthermore, the equality $(c)_{\mathfrak{B}} = \dfrac{(c)_{\mathfrak{B}\setminus\{d\}}}{1 - (d)_{\mathfrak{B}\setminus\{c\}}}$ holds for all neighbours c, d in $\mathfrak{B}$.

Proof. The numerator in formula (10) is just minus the sum of those terms in $[\mathfrak{B}]$ which have the factor c, and the denominator is the sum of those which do not have this factor. We prove (10) for connected diagrams $\mathfrak{B}$ and observe the general formula follows readily from this case. We fix a connected diagram $\mathfrak{B}$, a node c in $\mathfrak{B}$, and a neighbour d of c as in 9.4.(iii), and assume formula (10) holds for all connected (proper) subdiagrams of $\mathfrak{B}$. Then we prove it also holds for c and $\mathfrak{B}$.

We write $\mathfrak{B}\setminus\{c\}$ as a disjoint union $\mathfrak{C} \cup \mathfrak{D}$, where $\mathfrak{C}$ is the connected component of $\mathfrak{B}\setminus\{c\}$ containing d, and similarly put $\mathfrak{B}\setminus\{d\} = \mathfrak{E} \cup \mathfrak{F}$, where $\mathfrak{E}$ is the connected component of $\mathfrak{B}\setminus\{d\}$ which contains c. Since $\mathfrak{D} \cup \{c\}$ is connected and does not contain d, we have $\mathfrak{D} \cup \{c\} \subseteq \mathfrak{E}$. This implies the existence of a separation

$$\mathcal{E}\setminus\{c\} = \mathcal{D}\cup\mathcal{H},$$

where $\mathcal{H}$ is a union of connected components of $\mathcal{C}\setminus\{d\}$. The same argument shows that

$$\mathcal{C}\setminus\{d\} = \mathcal{F}\cup\mathcal{H},$$

and that this partition of $\mathcal{C}\setminus\{d\}$ is also a separation.

Using the induction hypothesis, we then have

$$(c)_{\mathcal{B}} = \frac{(c)_{\mathcal{E}}}{1-(d)_{\mathcal{C}}} = \frac{([\mathcal{E}\setminus\{c\}]-[\mathcal{E}])\,[\mathcal{C}\setminus\{d\}]}{[\mathcal{E}\setminus\{c\}]\,[\mathcal{C}]} = \frac{([\mathcal{E}\setminus\{c\}]-[\mathcal{E}])\,[\mathcal{F}]\,[\mathcal{H}]}{[\mathcal{D}]\,[\mathcal{H}]\,[\mathcal{C}]}$$

$$= \frac{[\mathcal{B}\setminus\{c,d\}]-[\mathcal{B}\setminus\{d\}]}{[\mathcal{B}\setminus\{c\}]} = \frac{[\mathcal{B}\setminus\{c\}]-[\mathcal{B}]}{[\mathcal{B}\setminus\{c\}]},$$

where the last equality holds because c and d are neighbours, and therefore cannot occur in the same non-zero term of $[\mathcal{B}]$. We have thus established formula (10). Using this formula, the proof of the last statement of the lemma is simply a verification - which would repeat the preceding calculation.□

Fix a vertex l in A and let $c_1, \dots, c_t$ be the simple cycles in A which contain l. We now use 9.5 to calculate the sum $(l) = \sum_{c\in\mathcal{A}(l)} wt_Q(c)$. We can write $(l) = (c_1)_{\mathcal{A}\setminus\{c_i : i\neq 1\}} + \dots + (c_t)_{\mathcal{A}\setminus\{c_i : i\neq t\}}$, and from here, using (10), it is a relatively simple exercise to conclude that

$$(l) = \frac{[\mathcal{A}\setminus\{c_1,\dots,c_t\}]-[\mathcal{A}]}{[\mathcal{A}\setminus\{c_1,\dots,c_t\}]} \qquad (11).$$

Formula (11) should be interpreted as follows: the left-hand side is a formal power series on the variables $c_1, \dots, c_k$, and the right-hand side is a rational function on the same variables; at the points $(c_1, \dots, c_k)$ $(c_i > 0)$ where the series (l) converges, its sum is given by the expression on the right-hand side. This is because the condition $(l) < \infty$, at some point $(c_1, \dots, c_k)$, ensures that all infinite series which occur in the proof converge at this point.

We now complete the proof of theorem 9.1. Assume 9.1.(ii) holds: then, for any vertex l, we have $(l) < \infty$ at the point $(wt_Q(c_1), \dots, wt_Q(c_k))$; and, using (11) and the assumption $[\mathcal{A}]$

$= 0$, we conclude that $(I) = 1$. We have proved 9.1.(ii)$\Rightarrow$9.1.(i).

Now assume 9.1.(i) holds. By 8.1 we know $(I) = 1 < \infty$ for all vertices I, and it follows from (11) that $[\mathcal{A}] = 0$. We have proved 9.1.(i) $\Rightarrow$ 9.1.(ii).$\square$

Remark 1. Assume again $(I) = 1$ for some (and therefore for every) vertex I. It is then easy to see that $(c)_{\mathfrak{B}} \leq 1$ for every subdiagram $\mathfrak{B}$ and $c \in \mathfrak{B}$, with equality iff $\mathfrak{B} = \mathcal{A}$. We therefore have, by formula (10), the inequalities

$$0 < \frac{[\mathfrak{B}\backslash\{c\}] - [\mathfrak{B}]}{[\mathfrak{B}\backslash\{c\}]} \leq 1,$$

with equality on the right iff $\mathfrak{B} = \mathcal{A}$. Assuming $[\mathfrak{B}\backslash\{c\}] > 0$, we obtain $[\mathfrak{B}\backslash\{c\}] > [\mathfrak{B}] \geq 0$ (with $[\mathfrak{B}] = 0$ iff $\mathfrak{B} = \mathcal{A}$). Since $[\emptyset] = 1$, this provides an inductive proof that $(I) = 1$ implies the following condition:

(α) $0 \leq [\mathfrak{B}] \leq 1$ for all subdiagrams $\mathfrak{B}$ of $\mathcal{A}$, with $[\mathfrak{B}] = 0$ iff $\mathfrak{B} = \mathcal{A}$, and $[\mathfrak{B}] = 1$ iff $\mathfrak{B} = \emptyset$.

Note that $[\mathfrak{B}\backslash\{c\}] - [\mathfrak{B}] = c[\mathfrak{B}\backslash(\{c\}\cup\{\text{neighbours of } c\})]$. Since the variables $c_1, \ldots, c_k$ are assumed to be positive, we conclude that condition (α) is equivalent to this other condition:

(β) $[\mathfrak{B}] \geq 0$, with equality iff $\mathfrak{B} = \mathcal{A}$.

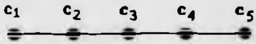
Assuming condition (β) holds we can prove (say, inductively) that $(c)_{\mathfrak{B}} \leq 1$ for every subdiagram $\mathfrak{B}$ and $c \in \mathfrak{B}$. Having thus concluded that $(I) < \infty$ we are free to use formula (11) to obtain $(I) = 1$.

Our next theorem, which is more convenient for applications than 9.1, summarises our conclusions:

9.6. Theorem. *Q has spectral radius 1 if and only if the conditions $[\mathfrak{B}] > 0$ (for any connected, proper subdiagram of $\mathcal{A}$) and $[\mathcal{A}] = 0$ hold for $(c_1, \ldots, c_k) =$*

$(wt_Q(c_1), \ldots, wt_Q(c_k))$.

Example. Consider the following loop diagram $\mathcal{A}$:



Then the equation $r_{\mathcal{A}} = 0$ can be given the form

$$\frac{c_3\,(1 - c_1)\,(1 - c_5)}{[1 - (c_1 + c_2)]\,[1 - (c_4 + c_5)]} = 1.$$

Putting $c_1 = c_4 = \frac{1}{2}$, $c_2 = c_5 = \frac{2}{3}$, we obtain $c_3 = \frac{1}{6}$. This solution $(\frac{1}{2}, \frac{2}{3}, \frac{1}{6}, \frac{1}{2}, \frac{2}{3})$ of the loop equation does not correspond to a matrix Q of spectral radius 1, because for instance $[\{c_1, c_2\}] = 1 - (c_1 + c_2) < 0$.

In all examples we had previously encountered (example 8.1 and the two examples in this section) a root $(c_1, \ldots, c_k)$ of the loop polynomial with $0 < c_i < 1$ would necessarily satisfy condition (β) above, and would therefore correspond by 9.6 to a matrix of spectral radius 1. This example is the simplest possible (e.g. with the fewest possible nodes) where this situation does not occur.

Remark 2. Karl Petersen [Pe] also applies loop methods to the computation of topological entropy, and his theorem 7.5 is similar to our lemma 8.1. The difference is that he does not restrict himself to finite-state chains (as we do), and we do not consider only measures of maximal entropy (as he does); also our approach to the proof is entirely unlike Petersen's. I thank Bill Parry for drawing my attention to Petersen's work.

Proof of 9.3. Before beginning the proof we introduce some notation. If $u_i = I\ U_i\ I$ $(i = 1, \ldots, r)$ are cycles based at the same vertex I then by $u_1 * u_2 * \ldots * u_r$ we denote the cycle $I\ U_1\ I\ U_2\ I \ldots I\ U_r\ I$. We shall use capital letters B, C to denote subgraphs of A, and letters $\mathcal{B}$, $\mathcal{C}$ to denote the corresponding subdiagrams of $\mathcal{A}$. If B is a connected subgraph of A and I, J are any two vertices in B then $\mathcal{B}(I)$ is the set of loops based at I which terminate at the second occurrence of I, $\mathcal{B}(I, J)$ is the set of loops in B that are based at I and visit J before the second occurrence of I, and $\mathcal{B}^*(I, J)$ is the subset of $\mathcal{B}(I, J)$ of the loops of the form $u * v$, where u and v are cycles based at I, u belongs to $\mathcal{B}(I)$, and J does not occur in v (we allow for v to be empty).

Our aim is to construct, for each pair of vertices I, J in $\mathcal{A}$, a bijection Φ_{IJ}: $\mathcal{A}(I, J) \to \mathcal{A}(J, I)$ with the following properties:

(i) $\Phi_{IJ}(u)$ has exactly the same factors as u, and occurring the same number of times;

(ii) if c is the first factor of u which contains J then c is the first factor of $\Phi_{IJ}(u)$.

We express (i) by saying Φ_{IJ} is a *reshuffling*, for that's how it acts on the level of factorizations. The existence of Φ_{IJ} suffices to prove the lemma, because from (i) and (ii) it follows that Φ_{IJ} preserves the weight and, if c is a simple cycle including I and J, that $\Phi_{IJ}(\mathcal{B}_c(I)) = \mathcal{B}_c(J)$ for every subdiagram $\mathcal{B}$ containing c.

The method of construction is inductive. We shall construct, for every connected subgraph B of A and every pair of vertices I, J in B, a bijection Φ^B_{IJ}: $\mathcal{B}(I, J) \to \mathcal{B}(J, I)$ with properties (i) and (ii) and also with the following additional properties:

(iii) $\Phi^B_{IJ}(\mathcal{B}^*(I, J)) \subset \mathcal{B}^*(J, I)$;

(iv) if $u = u_1 * \ldots * u_r$, where each u_i is in $\mathcal{B}^*(I, J)$, then $\Phi^B_{IJ}(u) = v_1 * \ldots * v_r$, where $v_i = \Phi^B_{IJ}(u_i)$ for each i.

The initial step in the inductive proccess is to consider subgraphs which consist of only one simple cycle, and for these subgraphs it is trivial to define the bijections. We now fix a connected subgraph B and two vertices I and J in B and, assuming we have already constructed Φ^C_{KL}: $\mathcal{C}(K, L) \to \mathcal{C}(K, L)$ with properties (i) - (iv) for all proper connected subgraphs C of B and all pairs of vertices K, L in C, we proceed to define Φ^B_{IJ}. In our construction it may sometimes happen that K = L and in that case Φ^C_{KL} is the identity. The following fact simplifies our task:

Claim 1. *If Φ^{B}_{IJ} is injective and has property (i) then Φ^{B}_{IJ} is a bijection.*

Proof. It follows from (i) that u and $\Phi^{B}_{IJ}(u)$ have the same length, and therefore all we have to prove is that $\mathfrak{B}(I, J)$ and $\mathfrak{B}(J, I)$ have the same number of cycles with length l, for any l. We do this by exhibiting a bijection Ψ: $\mathfrak{B}(I, J) \to \mathfrak{B}(J, I)$ that preserves the length.

Given $u \in \mathfrak{B}(I, J)$, we write $u = I\ T\ J\ U\ I\ V\ J\ W\ I$, where T, U, V, W are sequences of symbols, $I\ T\ J\ U\ I \in \mathfrak{B}(I)$, and J does not occur in neither U nor W. Then we define $\Psi(u) = J\ W\ I\ T\ J\ U\ I\ V\ J$ (if J does not occur in u after the second occurrence of I, we write $u = I\ T\ J\ U\ I\ W\ I$, with $I\ T\ J\ U\ I$, U and W as before, and let $\Psi(u) = J\ U\ I\ W\ I\ T\ J$). Given $v \in \mathfrak{B}(J, I)$, write $v = J\ W\ I\ T\ J\ \bar{U}\ I\ V\ J$, where $J\ W\ I\ \bar{T}\ J \in \mathfrak{B}(J)$ and I does not occur in neither T nor $\bar{U}$, and let $\Theta(v) = I\ \bar{T}\ J\ \bar{U}\ I\ V\ J\ W\ I$ (in the case where I does not occur in v after the second occurrence of J we write $v = J\ W\ I\ \bar{T}\ J\ \bar{U}\ J$ and let $\Theta(v) = I\ \bar{T}\ J\ \bar{U}\ J\ W\ I$). We see that if $v = \Psi(u)$ then we have, with the above notation, $I\ T\ J\ U\ I = I\ \bar{T}\ J\ U\ I$, and from here we conclude that $\Theta = \Psi^{-1}$.□

Now consider two cycles $u \in \mathfrak{B}(J) \cap \mathfrak{B}(J, I)$ and $v \in \mathfrak{B}(I) \setminus \mathfrak{B}(I, J)$ (i.e., v does not visit J). Write $u = J_0\ J_1 \ldots J_m$ with $J_0 = J_m = J$ and let $u_1, \ldots, u_r$ be the deleting loops of u numbered by their order of occurrence. Denote by c the first factor of u and write c $= J_{i_0} J_{i_1} \cdots J_{i_k}$, where $0 = i_0 < i_1 < \ldots < i_k = m$ and if $i_h > i_{h-1} + 1$ then the block $J_{i_{h-1}+1}\ J_{i_{h-1}+2} \cdots J_{i_h}$ has the form $u_n \bullet \ldots \bullet u_{n+j}$, for some consecutive deleting loops of u based at J_{i_h}. We now define the *inserting index* t(u, v) of v in u, which is a number between 0 and m. The definition is recursive, and we distinguish two cases:

(a) Some vertex of c occurs in v and, if J_{i_h} is the last vertex of c which occurs in v, we have $J_t \neq I$ for all $t > i_h$. In this case we define $t(u, v) = i_h$.

(b) The vertex I occurs in some deleting loop of u and, if $u_j = J_p \ldots J_{p+s}$ is the last deleting loop where I occurs, then J_{i_h} does not occur in v for all $i_h \geq p + s$. In this case I cannot be the base vertex of u_j. Write $u_j = j_0 \dots j_s$, where $j_0 = j_s$ and $j_i = j_{p+i}$ for all $0 \leq i \leq s$. Assuming we have defined the inserting index $t(u_j, v)$ - which makes sense because I occurs in u_j and the base vertex of u_j does not occur in v - we define $t(u, v) = t(u_j, v) + p$. Then $t(u, v)$ is the position in u corresponding to the position $t(u_j, v)$ in u_j.

The inserting index is also defined for $u \in \mathfrak{D}^*(J, I)$ and $v \in \mathfrak{D}(I)\setminus\mathfrak{D}(I, J)$: we write $u = \tilde{u}_1 * \tilde{u}_2$, where $u_1 \in \mathfrak{D}(J)$, and define $t(u, v) = t(\tilde{u}_1, v)$.

We notice the inserting index has the following property: the vertex I cannot occur in u after the position corresponding to the index $t(u, v)$. In case (a) this is obvious and otherwise it follows by induction. Now let $t = t(u, v)$ and $\Theta(v) = j_t V j_t$ be a cycle in $\mathfrak{D}^*(J_t, I)$ which has exactly the same vertices as v, and consider the cycle $w = j_0 j_1 \dots j_t V j_t j_{t+1} \dots j_m$ which we obtain by inserting $\Theta(v)$ in u at the position t. In case (a) $\Theta(v)$ is a deleting loop of w (or a block consisting of consecutive deleting loops with the same base vertex), and therefore the factorization of w is obtained by inserting the factorization of $\Theta(v)$ somewhere in the factorization of u. And this very important property of the factorizations is true also in case (b), as we easily check using induction.

Our next claim is the main burden of the proof, but it helps much in clarifying our construction.

Claim 2. *If there exists an injection Ψ: $\mathfrak{D}(I)\bigcap\mathfrak{D}(I, J) \to \mathfrak{D}(J, I)$ with properties* (i) *and* (ii) *then there also exists a bijection Φ^B_{IJ}: $\mathfrak{D}(I, J) \to \mathfrak{D}(J, I)$ with all the properties* (i) -(iv).

Proof. Since for $v \in \mathfrak{D}(I)\bigcap\mathfrak{D}(I, J)$ the vertex I occurs only once in $\Psi(v)$ (because it occurs in only one factor), we conclude that $\Psi(v) \in \mathfrak{D}^*(J, I)$. We define an injection Φ^B_{IJ}: $\mathfrak{D}^*(I, J) \to \mathfrak{D}^*(J, I)$ whose restriction to $\mathfrak{D}(I)\bigcap\mathfrak{D}(I, J)$ is equal to Ψ and then extend Φ^B_{IJ} to $\mathfrak{D}(I, J)$ by requiring it to have property (iv). This extension is again injective and has properties (i) - (iv), and therefore Φ^B_{IJ} is a bijection by claim 1.

Having defined Φ^B_{IJ} on $\mathfrak{D}(I)\bigcap\mathfrak{D}(I, J)$ as being equal to Ψ, we now define it on $\mathfrak{D}^*(I, J)$. If all cycles based at I visit J (i.e., if every connected subgraph of B which contains I also contains J) then we have $\mathfrak{B}^*(I, J) = \mathfrak{B}(I) \cap \mathfrak{B}(I, J)$ and we are done. Otherwise let C be the largest connected subgraph of B which contains I and *does not* contain J. For a cycle v in $\mathfrak{B}^*(I, J)$ we write it $v = v_1 * \ldots * v_r$, where each $v_i \in \mathfrak{B}(I)$. Then $v_1 \in \mathfrak{B}(I) \cap \mathfrak{B}(I, J)$ and each v_j ($j = 2, \ldots, r$) belongs to $\mathfrak{C}(I)$.

Let $u = \Psi(v_1)$ and write $u = J_0 J_1 \ldots J_m$. Roughly, we obtain $\Phi^B_{IJ}(v)$ by reshuffling the cycles $v_2, \ldots, v_r$ and inserting them in u. But first we define cycles w_i as follows: if $j_1 > 2$ is the first index such that $t(u, v_{j_1}) \geq t(u, v_2)$ we let $w_1 = v_2 * \ldots * v_{j_1-1}$; next, if $j_2 > j_1$ is the first index such that $t(u, v_{j_2}) \geq t(u, v_{j_1})$ we let $w_2 = v_{j_1} * v_{j_1+1} * \ldots * v_{j_2-1}$; we proceed in this way to define $w_3, \ldots, w_s$ so that $w_1 * \ldots * w_s = v_2 * \ldots * v_r$.

Write $t_i = t(u, w_i)$ for $i = 1, \ldots, s$, and notice that $t_1 \leq t_2 \leq \ldots \leq t_s$. Consider the map $\Phi^C_i: \mathfrak{C}(I, J_{t_i}) \to \mathfrak{C}(J_{t_i}, I)$, which we assume to have properties (i) - (iv), and let $\bar{w}_i = \Phi^C_i(w_i)$. Then each $\bar{w}_i = J_{t_i} W_i J_{t_i}$ belongs to $\mathfrak{C}^*(J_{t_i}, I)$, and we define $\Phi^B_{IJ}(v)$ as the cycle obtained by inserting the cycles $\bar{w}_i$ in u at the positions t_i. More precisely, we define

$$\Phi^B_{IJ}(u) = J_0 J_1 \ldots J_{t_1} W_1 J_{t_1} J_{t_1+1} \ldots J_{t_s} W_s J_{t_s} J_{t_s+1} \ldots J_{m-1} J_m,$$

with the obvious modifications when some $t_h = t_{h+1}$. It is clear that $\Phi^B_{IJ}(\mathfrak{B}^*(I, J)) \subseteq \mathfrak{B}^*(J, I)$, and it only remains to check that Φ^B_{IJ} is injective and has properties (i) and (ii).

Using the above notation, since the first factor of $\Phi^B_{IJ}(v)$ is equal to the first factor of $\Psi(v_1)$ and by assumption Ψ has property (ii), Φ^B_{IJ} also has property (ii). By the remarks following the definition of inserting index, the factorization of $\Phi^B_{IJ}(u)$ is obtained by inserting the factorizations of $\Phi^C_i(w_i)$ somewhere in the factorization of $\Psi(v_1)$; and, since Ψ and the Φ^C_i are reshufflings, it follows that Φ^B_{IJ} is also a reshuffling.

Now we prove Φ^B_{IJ} is injective. For a cycle $w \in \mathfrak{B}^*(J, I)$, we give an algorithm for deleting some loops in w. This algorithm, when applied to $\Phi^B_{IJ}(v)$, produces $u = \Psi(v_1)$, and the loops we delete are successively $\bar{w}_1, \ldots, \bar{w}_s$. The fact that Ψ and the Φ^C_i are injective then implies that Φ^B_{IJ} is also injective.

Step 0: u is equal to w;

$s := 0$;

Step 1: if I occurs only once in u then jump to step 5;

p is equal to the length of u;

the vertices $J_0, \dots, J_p$ are defined by $u = J_0 \dots J_p$;

w is equal to u;

$h := 0$;

Step 2: let m be the length of w and $L_j := J_{j+h}$ $(j = 0, \dots, m)$ so that $w = L_0 \dots L_m$;

let $L_t\, L_{t+1} \dots L_{t+r}$ be the deleting loop of w which contains the 2^{nd} occurrence of I in w (if there are two such loops then we mean the first one);

if $L_j = I$ for some $j \le t$ then jump to step 4, otherwise proceed to step 3;

Step 3: $h := h+t$ and $w := L_t\, L_{t+1} \dots L_{t+r}$;

return to step 2;

Step 4: $s := s + 1$;

$w_s := L_t\, L_{t+1} \dots L_{t+n}$, where $n \ge r$ is the largest integer such that $L_t\, L_{t+1} \dots L_{t+n}$ consists of consecutive deleting loops of w based at L_t and $L_t\, L_{t+1} \dots L_{t+n} \in \mathfrak{D}^*(L_t, I)$;

$u := J_0\, J_1 \dots J_{h+t-1}\, J_{h+t+n}\, J_{h+t+n+1} \dots J_p$;

return to step 1;

Step 5: end.

The output of the algorithm is the cycle u and, if $s \ge 1$, the cycles $\bar{w}_1, \dots, w_s$. By following the algorithm and comparing it with the definition of inserting index we check that, for $w = \Phi^B_{IJ}(v)$, these cycles are the ones we've defined before. This completes the proof that Φ^B_{IJ} is injective.□

Claim 2 reduces the proof of 9.3 to showing the existence of a *good injection* Ψ: $\mathfrak{D}(I) \cap \mathfrak{D}(I, J) \to \mathfrak{D}(J, I)$ (i.e., an injection with properties (i) and (ii)). We consider two cases:

(a) The vertex I has either two distinct predecessors or two distinct successors. Assume the first hypothesis holds (the other one is similar) and let K be one of the predecessors of I. Let C (respectively D) be the largest connected subgraph of B that contains I and does not contain the edge K I (resp. contains the edge K I and does not contain any other edge terminating at I).

If $v \in \mathfrak{B}(I) \cap \mathfrak{B}(I, J)$ then v belongs to exactly one of the sets $\mathfrak{C}(I)$ and $\mathfrak{D}(I)$; if it belongs to $\mathfrak{C}(I)$ define $\Psi(v) = \Phi^{C}_{IJ}(v)$ and if it belongs to $\mathfrak{D}(I)$ define $\Psi(v) = \Phi^{D}_{IJ}(v)$. This is injective because Φ^{C}_{IJ} and Φ^{D}_{IJ} are injective and because, say, the edge K I occurs in $\Psi(v)$ if and only if it occurs in v - i.e., if and only if $v \in \mathfrak{D}(I)$.

(b) The vertex I has only one successor and one predecessor. Let I_1 be the predecessor of I. If I is the only successor of I_1 then the sets $\mathfrak{B}(J, I)$ and $\mathfrak{B}(J, I_1)$ are identical and there exists an obvious bijection $\mathfrak{B}(I) \cap \mathfrak{B}(I, J) \to \mathfrak{B}(I_1) \cap \mathfrak{B}(I_1, J)$ that preserves factorizations. Therefore if there exists a good injection $\mathfrak{B}(I_1) \cap \mathfrak{B}(I_1, J) \to \mathfrak{B}(J, I_1)$ then there also exists a good injection $\mathfrak{B}(I) \cap \mathfrak{B}(I, J) \to \mathfrak{B}(J, I)$.

If I_1 has only one successor and one predecessor then we let I_2 be the predecessor of I_1 and repeat the above considerations. Ultimately we get to a vertex I_n with only one successor and such that the existence of a good injection $\mathfrak{B}(I_n) \cap \mathfrak{B}(I_n, J) \to \mathfrak{B}(J, I_n)$ is equivalent to the existence of a good injection $\mathfrak{B}(I) \cap \mathfrak{B}(I, J) \to \mathfrak{B}(J, I)$; and furthermore one of the following alternatives holds:

- I_n has at least two predecessors;
- I_n has only one predecessor K, and K has at least two successors.

The first alternative has been dealt with in case (a). Now we consider the second alternative, and write I instead of I_n. Since K has at least two successors, we have already shown (by (a) and claim 2) the existence of a bijection Φ^{B}_{KJ}: $\mathfrak{B}(K, J) \to \mathfrak{B}(J, K)$ with properties (i) - (iv), and we shall use Φ^{B}_{KJ} to construct Φ: $\mathfrak{B}(I) \cap \mathfrak{B}(I, J) \to \mathfrak{B}(J, I)$.

For $u \in \mathfrak{B}(I) \cap \mathfrak{B}(I, J)$, write $u = I\ W_1\ K\ W_2\ K \ldots K\ W_r\ K\ I$, where K does not occur in

W_1 and each K W_i K ($i = 2, \ldots, r$) is a cycle in $\mathfrak{B}(K)$, and let $\bar{u}$ = K I W_1 K W_2 K ... K W_r K. The cycles u and $\bar{u}$ have the same factorization and the correspondence $u \to \bar{u}$ is certainly injective, but perhaps $\bar{u}$ does not belong to $\mathfrak{B}(K, J)$. So we define a new cycle $\alpha(u)$ as follows:

- if $\bar{u} \in \mathfrak{B}(K, J)$ (i.e., if J occurs in W_1) then $\alpha(u) = \bar{u}$;

- if $\bar{u} \notin \mathfrak{B}(K, J)$ then let i be the minimum index such that J occurs in K W_i K, and define u_2 = K I W_1 K W_2 K ... K W_{i-1} K; let $j > i$ be the next index such that J occurs in I W_j I and define u_1 = K W_i K W_{i+1} K ... K W_{j-1} K and u_3 = K W_j K W_{j+1} K ... K W_r K (if j does not exist then u_3 is empty). We then have $\bar{u} = u_2 \bullet u_1 \bullet u_3$, and define $\alpha(u) = u_1 \bullet u_2 \bullet u_3$.

The proof that α is injective is straightforward (notice for instance that if K I is not the first edge of $\alpha(u)$ then u_2 is the largest block in $\alpha(u)$ which is a cycle based at K, begins with K I and does not have any occurrence of J). It is also clear that $\alpha(u) \in \mathfrak{B}(K, J)$, $u_1 \bullet u_2 \in \mathfrak{B}^*(K, J)$ and $u_3 \in \mathfrak{B}(K, J)$.

We define $\Psi(u) = \Phi^B_{KJ}(\alpha(u))$.

Write $\Psi(u) = v \bullet w$, where $v = \Phi^B_{KJ}(u_1 \bullet u_2)$ and $w = \Phi^B_{KJ}(u_3)$. Then $v \in \mathfrak{B}^*(J, K)$ and the vertex I occurs in v. Since every occurrence of I is immediatly preceded by an occurrence of K, we conclude that $v \in \mathfrak{B}^*(J, I)$, and therefore $\Psi(u) \in \mathfrak{B}(J, I)$. The fact that Ψ is a reshuffling and is injective follows from the same properties for α and Φ^B_{KJ}; and property (ii) is also easily checked. Therefore Ψ is a good injection and the proof of 9.3 is finished.□

Appendix.

We present here a generalization of proposition 5.2.

A.1. Proposition. *Within each non-trivial irreducible stochastic flow equivalence class there* *exists a matrix A with $\Gamma/\Delta \simeq \mathbb{Z}$.*

Proof. Given A, let $\exp a_1, \ldots, \exp a_s$ ($a_i > 0$) be a basis of Γ. For an edge e we put $wt_A(e) = \exp[-(a_1g_1(e) + \ldots + a_sg_s(e)]$, where the g_i are integer-valued functions, and define $\phi(e) = a_1g_1(e) + \ldots + a_sg_s(e)$. If $u = e_1 \ldots e_k$ is a path, let $\phi(u)$ denote the sum $\phi(e_1) + \ldots + \phi(e_k)$, and $g_i(u)$ the sum $g_i(e_1) + \ldots + g_i(e_k)$ ($1 \le i \le s$). Write

$$wt_A(u) = wt_1(u) \ldots wt_s(u),$$

where $wt_i(u) \in < \exp a_i >$. If we have $wt_1(u) < 1$, i.e. $g_1(u) > 0$, for all cycles u, we use the proof of 5.2 to construct a flow equivalent matrix for which a number $\gamma < 1$ exists such that $wt_1(u) = \gamma^{l(u)}$ for all cycles u. It then follows that $\Delta \subseteq <\exp a_2, \ldots, \exp a_s>$ and, since Γ/Δ is cyclic, that $\Gamma/\Delta \simeq \mathbb{Z}$.

Otherwise, since we have $\phi(u) = a_1g_1(u) + \ldots + a_sg_s(u) > 0$ for all simple cycles u, we may choose relatively prime integers $m_1, \ldots, m_s$ such that $m_1g_1(u) + \ldots + m_sg_s(u) > 0$ for all simple cycles u and therefore for all cycles. By theorem II.1 in [N], we choose an integer matrix $\mathcal{A} = (\alpha_{ij})$ with determinant equal to 1 and such that the first row of $\mathcal{A}^{-1}$ is $(m_1, \ldots, m_s)$. Set

$$\mathbf{a} = [\, a_1 \quad \ldots \quad a_s \,],\ \mathbf{g} = \begin{bmatrix} g_1 \\ \ldots \\ g_s \end{bmatrix}.$$

Then $\exp(\mathbf{a}\mathcal{A})_1, \ldots, \exp(\mathbf{a}\mathcal{A})_s$ form a basis of Γ, and we have the equality

$$\phi = \mathbf{a}\mathbf{g} = (\mathbf{a}\mathcal{A})(\mathcal{A}^{-1}\mathbf{g})$$
$$= (\mathbf{a}\mathcal{A})_1(\mathcal{A}^{-1}\mathbf{g})_1 + \ldots + (\mathbf{a}\mathcal{A})_s(\mathcal{A}^{-1}\mathbf{g})_s.$$

Now we know that $(\mathcal{A}^{-1}\mathbf{g})_1(u) = m_1g_1(u) + \ldots + m_sg_s(u)$ is strictly positive for all cycles u. We can repeat the above reasoning to justify the existence of a flow equivalent matrix with $\Delta \subseteq <\exp(\mathbf{a}\mathcal{A})_2, \ldots, \exp(\mathbf{a}\mathcal{A})_s >$ and $\Gamma/\Delta \simeq \mathbb{Z}$.$\square$

Almost flow equivalence for subshifts of finite type with finite group actions

10. *Finite group actions.*

We now consider the problems of flow equivalence and almost flow equivalence of irreducible SFT when there is a finite group action involved. To prepare the introduction of these concepts we have to say what sort of group actions we allow for on SFT and its suspensions and how the factor maps are to reflect this additional structure.

Throughout this chapter, G denotes a finite group. A *G-action* on an SFT Σ_A is an isomorphism of G into the subgroup of automorphisms of Σ_A which commute with the shift. A *special G-action* has the property that $gx \neq x$ for every x if $g \neq id$. Henceforth all G-actions are assumed to be special, and the term G-action is an abbreviation for special G-action.

By (Σ_A, G) we denote an SFT Σ_A with some fixed G-action. We say that (Σ_A, G) and (Σ_B, G) are conjugate if there exists a topological conjugacy π: $\Sigma_A \to \Sigma_B$ which preserves (or commutes with) the G-actions; and π: $(\Sigma_A, G) \to (\Sigma_B, G)$ is a factor map if π commutes with the G-actions, besides being a factor map as we have defined them between SFT.

Skew-products are a natural source for G-actions: if ϕ: $\Sigma_C \to G$ is some function then $\Sigma_C \times G$ is an SFT with the shift map defined by $\sigma_\phi(x,h) = (\sigma x, h\phi(x))$; the G-action is defined by $g(x,h) = (x,gh)$. It is known (and quite easy to prove) that any (Σ_A, G) is conjugate to some skew-product. But now we choose rather to look at G-actions from an intrinsic point of view.

Each (Σ_A, G) is conjugate to a one-block action, i.e. to a pair (Σ_B, G) where each $g \in G$ is a one-block map: $g(x_n)_{n\in\mathbb{Z}} = (gx_n)_{n\in\mathbb{Z}}$. This is proved in 2.(i) [AKM]. We shall always assume this recoding has been performed on any given (Σ_A, G), so that all G-actions are one-block actions. Since we want to consider matrices with non-negative (not just 0-1) matrices, it is useful to visualize the G-action as a group of automorphisms of the graph A as follows: each g defines a permutation of the vertices and, for each pair of vertices I and J, a bijection from the set $\mathcal{E}(I, J) = \{e: e$ is and edge from I to J$\}$ onto $\mathcal{E}(gI, gJ)$. Since each $g \neq$ id does not have any fixed points as an automorphism of the shift space, we may assume (by going to higher blocks) that it also does not fix any vertex as an automorphism of the graph.

Now we pass to G-actions on flow spaces $(\Omega, (\varphi_t)_{t\in\mathbf{R}})$. Since actual parametrizations are not important, we shall only require that each g is a homeomorphism that sends orbits onto orbits (preserving their orientations) and (if $g \neq$ id) has no fixed points. For a G-action (Σ_A, G) and a function $\alpha: \Sigma_A \to \mathbf{R}^+$ there exists a natural G-action on the suspension space $(\Sigma_A^\alpha, (\sigma_A^t)_{t\in\mathbf{R}})$ defined by $g[x,t] = [gx, t\alpha(gx)/\alpha(x)]$. Then $\Sigma_A \doteq \Sigma_A \times \{0\}$ is a (global) cross-section of $(\Sigma_A^\alpha, (\sigma_A^t)_{t\in\mathbf{R}}, G)$ that is invariant under the G-action.

These remarks motivate the following definition: (Σ_A, G) and (Σ_B, G) are *G-flow equivalent* (or simply flow equivalent) if there exists $(\Omega, (\varphi_t)_{t\in\mathbf{R}}, G)$ such that both Σ_A and Σ_B embed in Ω as G-invariant global cross-sections (on each section the shift being the Poincaré map and the G-action the restriction of the G-action on Ω). The next result is reassuring:

10.1. Proposition. *If Σ is a G-invariant global cross-section of $(\Omega, (\varphi_t)_{t\in\mathbf{R}}, G)$ and $\sigma: \Sigma \to \Sigma$ is the Poincaré map then the induced G-action on Σ commutes with σ.*

Proof. Let $\alpha: \Sigma \to \mathbf{R}^+$ be the first return time and, for $x \in \Sigma$, let $S_x = \{\varphi_t x: 0 \leq t \leq \alpha(x)\}$ be the portion of the orbit of x between x and $\sigma(x)$. The intersection $S_x \cap \Sigma$ consists of only the points x and σx, and therefore $g(S_x \cap \Sigma) = (gS_x) \cap \Sigma$ contains only the points gx and $g\sigma x$. This means that $g\sigma x$ is the first intersection of the positive orbit of gx with Σ, which is to say that $g\sigma x = \sigma g x$.□

By $(\Sigma_A^\alpha, (\sigma_A^t)_{t\in\mathbf{R}}, G)$ (or simply (Σ_A^α, G)) we denote some fixed G-action on Σ_A^α that leaves the base space $\Sigma_A \times \{0\}$ invariant. If (Σ_A, G) and (Σ_B, G) are flow equivalent then there are suspensions and G-actions (Σ_A^α, G) and (Σ_B^β, G) which are conjugate by a flow-commuting, G-commuting homeomorphism. Here the G-actions on the suspension spaces have only to respect the existing actions on the base spaces $\Sigma_A \times \{0\}$ and $\Sigma_B \times \{0\}$. By reparametrization we obtain a homeomorphism π: $(\Sigma_A^1, G) \to (\Sigma_B^1, G)$ that satisfies the requirements defining a factor map between continuous flows (section 3, conditions (i), (ii), (iv) and (v)) and also the following conditions:

(vi) $\pi^{-1}(\Sigma_B \times \{0\})$ is a G-invariant subset of Σ_A^1;

(vii) $\pi\, g\, O(\hat{x}) = g\, \pi\, O(\hat{x})$ for every $\hat{x}$ in Σ_A and g in G.

By $O(x)$ we denote the orbit of $\hat{x}$; note that $gO(x) = O(gx)$ and $\pi O(\hat{x}) = O(\pi x)$. Condition (vii) says that π and g commute as maps of the spaces of orbits. In this context, a factor map π: $(\Sigma_A^1, G) \to (\Sigma_B^1, G)$ is understood to be a map satisfying conditions (vi), (vii) and 3.(i), (ii), (iv), (v). Generalizing the proof of proposition 3.1, we find in the next proposition that any factor map can be replaced by a map with much better commuting properties. It follows in particular that (Σ_A, G) and (Σ_B, G) are flow equivalent if and only if there exists a homeomorphism $(\Sigma_A^1, G) \to (\Sigma_B^1, G)$ that is also a factor map.

10.2. Proposition. *Let π: $(\Sigma_A^1, G) \to (\Sigma_B^1, G)$ be a factor map, where A and B are irreducible matrices. Then there are a function α: $\Sigma_A \to \mathbf{Q}^+$ depending on finitely many coordinates; G-actions on Σ_A^α and on Σ_B^1 which commute with the flow and induce in the base spaces the same actions as before; and a factor map π: $(\Sigma_A^\alpha, G) \to (\Sigma_B^1, G)$ for these new actions which commutes with both the flow and the G-action.*

Remark. The maps π and π establish the same correspondence of orbits, and therefore π is as many to one (e.g. finite to one, almost everywhere one to one, homeomorphism) as π.

Proof. The proof of 3.1 gives us a flow-commuting factor map π: $\Sigma_A^{\hat{\alpha}} \to \Sigma_B^1$, where $\hat{\alpha}$: $\Sigma_A \to \mathbf{Q}^+$ depends only on finitely many coordinates. We now want to replace $\hat{\alpha}$ by a cohomologous function α such that $\alpha \circ g = \alpha$ for every g in G. For such a function α to exist we must have, for every periodic point x,

$$\tilde{\alpha}(x) + \ldots + \tilde{\alpha}(\sigma^{p-1}x) = \tilde{\alpha}(gx) + \ldots + \tilde{\alpha}(\sigma^{p-1}gx) \qquad (12),$$

where p is the period of x. A way of restating this is that for every periodic orbit τ in $\Sigma_A^{\tilde{\alpha}}$ the lengths of τ and $g\tau$ should be equal, and this condition does indeed hold:

10.3. Lemma. *If τ is a periodic orbit in $\Sigma_A^{\tilde{\alpha}}$ and $g \in G$ then $l(\tau) = l(g\tau)$.*

Proof. If η is a periodic orbit in Σ_B^1 we easily check that $l(\eta) = l(g\eta)$. The result follows for τ by letting $\eta = \tilde{\pi}(\tau)$ and noticing that, as a consequence of condition (vi), if $\tilde{\pi}_{|\tau}: \tau \to \eta$ is n to one then $\tilde{\pi}_{|g\tau}: g\tau \to g\eta$ is also n to one: then, since $\tilde{\pi}$ commutes with the flow, we have $l(\tau) = n\, l(\eta)$ and $l(g\tau) = n\, l(g\eta)$.□

Thus $\tilde{\alpha}$ satisfies the necessary condition (12) for the function α to exist. Happily this condition is also sufficient: we just define α by $\alpha(x) = \frac{1}{|G|}\sum_{g\in G}\tilde{\alpha}(gx)$. By (12) we conclude that $\alpha(x) + \ldots + \alpha(\sigma^{p-1}x) = \tilde{\alpha}(x) + \ldots + \tilde{\alpha}(\sigma^{p-1}x)$ whenever x is a periodic point of period p, and it is a standard result (see for instance [PP]) that this implies α and $\tilde{\alpha}$ are cohomologous with the cobounding function locally constant (because α and $\tilde{\alpha}$ are also locally constant).

Now that we have a flow-commuting factor map π: $\Sigma_A^{\alpha} \to \Sigma_B^1$ we will modify it to obtain another flow-commuting factor map $\hat{\pi}$ satisfying the following condition:

- for every x in Σ_A, if $\hat{\pi}[x, 0] = [y, s]$ then $\hat{\pi}[gx, 0] = [gy, s]$ (13).

We define $\hat{\pi}$ on $\Sigma_A \times \{0\}$ and extend to Σ_A^{α} by using the flow. If $x \in \Sigma_A$ is such that $\hat{\pi}O([x,0])$ is not a periodic orbit then, for $[y,0] \in \pi O([x,0])$, there is a unique number $s(x, y)$ with $\pi[x,0] = [y, s(x, y)]$. We define $\hat{s}(x, y) = \frac{1}{|G|}\sum_{g\in G} s(gx, gy)$, and list the following easy properties of $\hat{s}$:

(a) $\hat{s}(gx, gy) = \hat{s}(x, y)$ for every $g \in G$;

(b) $\hat{s}(x, \sigma^k y) = \hat{s}(x, y) - k$ for every $k \in \mathbb{Z}$;

(c) $\bar{s}(\sigma x, y) = \bar{s}(x, y) + \alpha(x)$.

Equality (c) follows from the similar equality for s, which holds because π preserves the parametrization of the orbits, and from the fact that $\alpha \circ g = \alpha$. If the orbit $\pi O([x,0])$ is not periodic and $[y,0]$ is a point in it, we set $\bar{\pi}[x,0] = [y, \bar{s}(x, y)]$: by (b) this definition does not depend on the choice of y; by (a) $\bar{\pi}$ satisfies condition (13); and by (c) we may extend $\bar{\pi}$ by letting $\bar{\pi}[x, t] = [y, \bar{s}(x, y) + t]$ for every $t \in \mathbb{R}$.

Let S be the set of the points $\hat{x}$ in Σ_A^{α} such that the orbit $\pi O(\hat{x})$ is not periodic. Then $\bar{\pi}$ so far has only been defined on S, which is a dense subset of Σ_A^{α} - but, since π and g are uniformly continuous (being continuous functions on compact metric spaces), $\bar{\pi}$ is uniformly continuous and we can extend it continuously to the whole of Σ_A^{α}. By continuity we have $\bar{\pi}O(x) = \pi O(\hat{x})$ for every point x in Σ_A^{α}; and, again by continuity, $\bar{\pi}$ satisfies condition (13).

Now we define new G-actions on Σ_A^{α} and Σ_B^{1}, denoting them by $\bar{G}$. The compatibility conditions which have to be verified are that $\bar{G}_{|\Sigma_A \times \{0\}} = G_{|\Sigma_A \times \{0\}}$ and $\bar{G}_{|\Sigma_B \times \{0\}} = G_{|\Sigma_B \times \{0\}}$. We simply define, for $g \in G$, $\bar{g}[x,t] = [gx,t]$. These new actions commute with the flow; condition (13) and the fact that π also commutes with the flow ensure that $\bar{\pi}\bar{g} = \bar{g}\bar{\pi}$ for every g.□

Again assume we have (Σ_A, G) with Σ_A irreducible. Let β: $\Sigma_A \to \mathbb{N}$ be a continuous function such that $\beta \circ g = \beta$ for every g, and let $\Sigma_{A_\beta} = \{(x, i) \in \Sigma_A \times \mathbb{N}: 1 \leq i \leq \beta(x)\}$ be the SFT with shift map defined by $\sigma_{A_\beta}(x, i) = (x, i+1)$ if $i < \beta(x)$, $\sigma_{A_\beta}(x, \beta(x)) = (\sigma x, 1)$. The G-action on Σ_{A_β} is defined by $g(x, i) = (gx, i)$. We check that (Σ_A, G) and (Σ_{A_β}, G) are flow equivalent by noticing there is an easy identification of the suspension spaces Σ_A^{β} and $\Sigma_{A_\beta}^{1}$, and this identification respects the natural G-actions. We say that (Σ_{A_β}, G) was obtained from (Σ_A, G) by *expansion*.

Using the techniques in the proof of 3.1 we establish the following consequence of 10.2:

10.4. Corollary. *Let A and B be irreducible matrices and* π: $(\Sigma_A^1, G) \to (\Sigma_B^1, G)$ *a factor* *map. Then there exists a factor map* $\pi_0: (\Sigma_{A_{\bar\beta}}, G) \to (\Sigma_{B_n}, G)$, *where* β *is some integer-valued positive continuous function such that* $\beta \circ g = \beta$ *for every* g, *and* n *is some positive integer. The map* π_0 *is finite to one (respectively a.e. one to one, homeomorphism) if and only if* π *is finite to one (respectively a.e. one to one, homeomorphism).*

We also have the following version of 3.2:

10.5. Corollary. *G-flow equivalence (between G-actions on irreducible SFT) is generated by conjugacy and expansion.*

Remark 1. For the sake of completeness, we prove here that every $(\Omega, (\varphi_t)_{t\in\mathbf{R}}, G)$ (one-dimensional hyperbolic flow with some G-action on it) possesses some G-invariant global cross-section.

We start with a global cross-section Σ and prove that, by slightly modifying it, we can construct another cross-section Σ such that the set $g\Sigma\bigcap\Sigma$ is empty for all $g \in G$. This implies that the distance (along the flow) between the different cross-sections $g\Sigma$ $(g \in G)$ is positive, and therefore $\bigcup_{g\in G} g\Sigma$ is again a cross-section which is obviously G-invariant.

Since Σ is compact and zero-dimensional and $gx \neq x$ for all g and x, we can construct an open-closed partition $U_1, \dots, U_r$ of Σ such that, for some $\epsilon > 0$ and all $1 \le i \le r$, the set $gU_i^\epsilon \bigcap U_i^\epsilon$ is empty, where $U_i^\epsilon = \bigcup_{-\epsilon\le t\le\epsilon} \varphi_t U_i$. Now we construct $V_1, \dots, V_r$ inductively as follows. We let $V_1 = U_1$. Then we find an open-closed partition $U_{21}, \dots, U_{2k}$ of U_2 and numbers $t_{21}, \dots, t_{2k}$ such that $|t_{2i}| \le \epsilon$ and $\varphi_{t_{2i}} U_{2i}\bigcap gV_1 = \emptyset$ for all g and i, and let $V_2 = \bigcup_{j=1}^{k} \varphi_{t_{2j}} U_{2j}$. Similarly we define a partition $U_{31}, \dots, U_{3r}$ of U_3 and numbers $t_{31}, \dots, t_{3r}$, this time taking care that $\varphi_{t_{3i}} U_{3i} \bigcap g(V_1 \bigcup V_2) = \emptyset$, and let $V_3 = \bigcup_{j=1}^{r} \varphi_{t_{3j}} U_{3j}$. We proceed in this way to define the V_i up to V_n, thus ensuring that $V_i \bigcap gV_j$ is empty for all $g \in G$ and i, j, and then let $\Sigma = \bigcup_{i=1}^{n} V_i$.

Remark 2. The definition of factor map π: $(\Sigma^1_A, G) \to (\Sigma^1_B, G)$ we gave is unsatisfactory in that the composition of factor maps is not necessarily a factor map. Here we propose an alternative, less restrictive definition that does not have this drawback and for which Proposition 10.2 remains true. We replace conditions (vi) and (vii) by the following single condition:

(vi') For each $g \in G$ there exists a continuous function $\rho(\ .\ , g)$: $\Sigma^1_A \to \mathbf{R}$ such that $\pi g(x) = \sigma_B^{\rho(x,\, g)} g\pi(x)$ for all $x \in \Sigma^1_A$.

We check that (vi') is preserved under composition of maps (the last paragraph of page 13 is helpful here). The fact that the former conditions imply (vi') is then a consequence of Proposition 10.2. For it follows that a map π: $(\Sigma^1_A, G) \to (\Sigma^1_B, G)$ satisfying those conditions can be written as a composition $(\Sigma^1_A, G) \to (\Sigma^\alpha_A, \tilde{G}) \xrightarrow{\tilde{\pi}} (\Sigma^1_A, G) \to (\Sigma^1_B, G)$, where all the maps satisfy (vi').

Let us agree that henceforth factor map means a map satisfying conditions (vi') and 3.(i), (ii), (iv), (v). We now prove that 10.2 and consequently 10.4 and 10.5 still hold true for this new definition. A glance at the proof of 10.2 convinces us it is enough to show lemma 10.3 remains valid, and this is a consequence of the following claim:

10.6. Claim. *Let π: $(\Sigma^1_A, G) \to (\Sigma^1_B, G)$ be a factor map and let η, τ be periodic orbits such that $\pi(\eta) = \tau$. If $\pi|_\eta$: $\eta \to \tau$ is n to one then $\pi|_{g\eta}$: $g\eta \to g\tau$ is also n to one.*

Proof. Choose $x \in \eta$ and let $y = \pi x$. Let m, $\bar{m}$, r, $\bar{r}$ be respectively the minimum positive periods of the orbits η, $g\eta$, τ, $g\tau$. Then there are increasing homemorphisms

$$[0, m] \to [0, \bar{m}] \qquad [0, m] \to [0, nr] \qquad [0, nr] \to [0, n\bar{r}]$$
$$t \to \bar{t} \qquad\qquad t \to s \qquad\qquad s \to \bar{s}$$

such that the following equations hold: $g\ \sigma^t_A\ (x) = \sigma^{\bar{t}}_A\ g\ (x)$, $\pi\ \sigma^t_A\ (x) = \sigma^s_B(y)$, $g\ \sigma^s_B\ (y) = \sigma^{\bar{s}}_B\ g\ (y)$. By (vi'), we can write

$$\pi\, g\, (\sigma_A^t x) = \sigma_B^{\rho(\sigma_A^t x,\, g)} g \pi(\sigma_A^t x),$$

for all $t \in [0, m]$. Using the equations above, we obtain

$$\pi\, \sigma_A^{\bar t}\, g\, (x) = \sigma_B^{\rho(\sigma_A^t x,\, g)} g\, \sigma_B^{s}(y)$$

$$= \sigma_B^{\rho(\sigma_A^t x,\, g) + \bar s} g(y).$$

Consider the function $\bar t \rightarrow \rho(\sigma_A^t x,\, g) + \bar s$, defined on $[0, m]$, and write $f_1(\bar t) = \rho(\sigma_A^t x,\, g)$, $f_2(\bar t) = \bar s$. The function f_2 has image $[0, nr]$ and f_1 is such that $f_1(0) = f_1(m)$. Therefore the image of the sum $f_1 + f_2$ contains an interval of length nr, and we conclude that $\pi|_{g\eta}$ is at least n to one, and by applying the same reasoning to g^{-1} we see that $\pi|_{g\eta}$ is exactly n to one.□

11. *Almost G-flow equivalence.*

We shall say $(\Sigma_A,\, G)$ and $(\Sigma_B,\, G)$ are *almost G-flow equivalent* if there exist $(\Sigma_C^1,\, G)$ and G-actions on Σ_A^1 and Σ_B^1 compatible with the existing actions on the base spaces such that both $(\Sigma_A^1,\, G)$ and $(\Sigma_B^1,\, G)$ are a.e. one to one factors of $(\Sigma_C^1,\, G)$. Our next theorem generalizes 6.1, and shows that G-actions on suspension spaces of SFT with positive entropy are indistinguishable from the point of view of almost G-flow equivalence.

11.1. Theorem. *Any two G-actions on irreducible SFT with positive entropy are almost flow equivalent.*

The proof is an elaboration of our proof of 6.1 and is based on the theorem of Adler, Kitchen & Marcus [AKM] which says that if Σ_A and Σ_B are aperiodic and have the same topological entropy then $(\Sigma_A,\, G)$ and $(\Sigma_B,\, G)$ are almost conjugate for any two G-actions on Σ_A and Σ_B (i.e., there exists $(\Sigma_C,\, G)$ such that both $(\Sigma_A,\, G)$ and $(\Sigma_B,\, G)$ are a.e. one to one factors of $(\Sigma_C,\, G)$). Theorem 11.1 is then an immediate consequence of our next proposition, which refines and generalizes 6.2.

11.2. Proposition. *Let G be a finite group acting on the irreducible SFT Σ_A. If Σ_A has positive entropy then there exists an aperiodic SFT Σ_B with entropy log 2 and a G-action on Σ_B such that (Σ_A, G) and (Σ_B, G) are flow equivalent.*

Proof of proposition 11.2. As usual we assume each $g \in G$ is a 1-block map, i.e. we have $g(x_n)_{n\in\mathbb{Z}} = (gx_n)_{n\in\mathbb{Z}}$, and that each $g \neq \mathrm{id}$, as a graph automorphism, does not fix any vertex. We use the notation Out(I) for the number of edges with initial vertex I.

First we show (Σ_A, G) is flow equivalent to $(\Sigma_{\tilde{A}}, G)$, where $\tilde{A}$ is a matrix such that $\mathrm{Out}(J) \geq 2$ for every vertex J and, for some vertex I, $\mathrm{Out}(I) \geq 3$. This is done in two steps: the first consists of creating I with $\mathrm{Out}(I) \geq 3$, the second of getting rid of the vertices J with $\mathrm{Out}(J) = 1$. In the final step of the proof we create two cycles u and v of equal lengths and such that the cycles gu and u (and also gv and v) do not overlap if $g \neq \mathrm{id}$, and hu and v do not overlap for any h; we then increase the length of u by 1 and keep the length of v unchanged, thus creating an aperiodic matrix which we show to have entropy log 2. The techniques we use are fairly standard, and the only additional complication is that the G-action has to be respected at each step.

1[st] step. *(Σ_A, G) is conjugate to $(\Sigma_{\tilde{A}}, G)$, where $\tilde{A}$ has some vertex I with $\mathrm{Out}(I) \geq 3$.*

If Σ_A has positive entropy and $\mathrm{Out}(J) \leq 2$ for every vertex J, we choose a vertex I with $\mathrm{Out}(I) = 2$. Going to higher blocks if necessary, we may assume that gI is not a follower of I if $g \neq \mathrm{id}$: if for instance the edges with initial vertex I are f_1 and f_2, connecting I to respectively gI and hI $(g \neq h)$, then the followers of the block $U = (h^{-1}g^{-1}f_1)\,(g^{-1}f_1)$ are $(g^{-1}f_1)\,f_1$ and $(g^{-1}f_1)\,f_2$, neither of which is gU or hU (if $g = h$ then we take $U = (g^{-2}f_2)\,(g^{-1}f_1)$). Next, again by going to higher blocks, we ensure the existence of a vertex $J \notin GI = \{gI\colon g \in G\}$ with $\mathrm{Out}(J) = 2$. Then we consider a simple path $e_1 \ldots e_t$ which has initial vertex I and final vertex gJ, for some $g \in G$, and does not visit any other vertices in the set $G\{I, J\}$. Call I_i to the final vertex of the edge e_i, and let $r \geq 1$ be the minimum index such that $Out(I_r) = 2$.

We shall concentrate on the path $e_1 \ldots e_r$. Let $I_0 = I$. Then we have $Out(I_0) = Out(I_r) = 2$ and $Out(I_i) = 1$ for $0 < i < r$, and therefore $gI_i \neq I_r$ for all g and $0 < i < r$. By construction $gI_0 \neq I_i$ for $1 \leq i \leq r$. If $0 < i < j < r$ then any path with initial vertex I_i starts with $e_{i+1} \ldots e_j$, which connects I_i to I_j, and therefore any path beginning at gI_i starts with $(ge_{i+1}) \ldots (ge_j)$, which connects gI_i to gI_j, and this implies we cannot have $gI_i = I_j$.

Now we split the vertex I_r by successors. Let f_1 and f_2 be the edges with initial vertex I_r. We define a new matrix C and G-action by the following data: the set of vertices is the union of $V_A \backslash GI_r$ with the sets Gf_1 and Gf_2; if $J, K \notin GI_r$ then $C(J,K) = A(J,K)$; if $J \notin GI_r$, $i = 1$ or $i = 2$, then $C(J,gf_i) = A(J,gI_r)$; if $J_i \notin GI_r$ is the final vertex of f_i $(i=1, 2)$, then there is one edge connecting gf_i to gJ_i; if the final vertex of f_i is hI_r, then for each g there is one edge connecting gf_i to ghf_j, for $j = 1, 2$.

The G-action on Σ_C is defined in an obvious way, and an easy argument shows that this new SFT is conjugate to Σ_A by a conjugacy which preserves the group action. If $r > 1$ then we consider the path $e_1 \ldots e_{r-1}$ in the new graph, which has only been affected by the state-splitting in that its final vertex I_{r-1} has now the two followers f_1 and f_2, and repeat the above process r-1 times. When we finish the process the vertex I_0 will have $Out(I_0) \geq 3$, and $\tilde{A}$ is the matrix we finally obtain.

2^{nd} step. *If A has a vertex I with $Out(I) \geq 3$ then there exists a matrix $\tilde{A}$ where $Out(J) \geq 2$ for every vertex J and where we still have $Out(I) \geq 3$, and such that $(\Sigma_{\tilde{A}}, G)$ is flow equivalent to (Σ_A, G).*

Assume J is a vertex in A with $Out(J) = 1$, and let JK be the one edge with initial vertex J. We remark that K cannot be of the form gJ, for otherwise an easy argument shows that all paths from J would wind around the cycle $J(gJ)(g^2J) \ldots (g^{k-1}J)J$, where k is the order of g. We show that the matrix C obtained by deleting the states GJ and replacing the edges with final vertex gJ by edges with final vertex gK is flow equivalent to A. By repeating this operation enough times we produce a matrix $\hat{A}$ where $\mathrm{Out}(J) \geq 2$ for every vertex J, and where the vertex I with $\mathrm{Out}(I) \geq 3$ was not destroyed.

Now we describe C: the set of vertices is $V_C = V_A \setminus GJ$; if L, M $\in V_C$ and M is not of the form gK then $C(L,M) = A(L,M)$; if $L \in V_C$ and $g \in G$ then $C(L,gK) = A(L,gK) + A(L,gJ)$.

To show C is flow equivalent to A we describe the operations which lead from A to C and preserve the flow equivalence class. The first operation is to split K by predecessors. We let $\mathcal{E}(K)$ be the set of *edges* with final vertex K and $\mathcal{E}^*(K) = \mathcal{E}(K)\setminus\{JK\}$. Since $gK \neq K$ for $g \neq id$, the edge g(JK) does not belong to $\mathcal{E}^*(K)$ for any $g \in G$. We define a new matrix D as follows:

- V_D is the union of $V_A \setminus GK$ with the sets $G\mathcal{E}^*(K)$ and $G(JK)$;
- if L, M $\in V_A \setminus GK$ then $D(L,M) = A(L,M)$;
- if $L \in V_A \setminus GK$ and $g \in G$ then $D(g\mathcal{E}^*(K),L) = D(g(JK),L) = A(gK,L)$;
- if $g \in G$ then $D(gJ,g(JK)) = A(gJ,gK) = 1$;
- if $L \in V_A \setminus GK$ is the initial vertex of some edge in $g\mathcal{E}^*(K)$ then $D(L,g\mathcal{E}^*(K)) =$ $A(L,gK)$;
- if g, h $\in G$ then $D(h\mathcal{E}^*(K),g\mathcal{E}^*(K)) = D(h(JK),g\mathcal{E}^*(K)) = A(hK,gK)$.

The matrix D is conjugate to A and the splitting was done so that the G-action was respected. We notice that for each $g \in G$ the vertex gJ is the only predecessor of g(JK), and g(JK) is the only follower of gJ. Therefore the matrix E, which is obtained from D by deleting the vertices of the form gJ and setting $E(L,M) = D(L,M)$, $E(g(JK),M) = D(g(JK),M)$, $E(L,g(JK)) = D(L,gJ)$ and $E(g(JK),h(JK)) = D(g(JK),hJ)$ for $L,M \in V_D \setminus G\{J, JK\}$ and $g,h \in G$, is flow equivalent to D.

Now we look at the matrix E: the vertices $g\mathcal{E}^*(K)$ and g(JK) have exactly the same followers, i.e. $E(g\mathcal{E}^*(K),L) = E(g(JK),L)$ for all vertices L. Therefore we can amalgamate them: the matrix C is obtained by replacing each pair of vertices $g\mathcal{E}^*(K)$ and g(JK) by the one vertex gK; if L,M $\notin$ GK then $C(L,M) = E(L,M)$, $C(gK,M) = E(g(JK),M)$ and $C(L,g(K)) = E(L,g(JK)) + E(L,g\mathcal{E}^*(K))$; if g,h $\in G$ then $C(gK,hK) = E(g(JK),h(JK)) + E(g(JK),h\mathcal{E}^*(K))$. This matrix is conjugate to E and the reader may check this is exactly the matrix C we have defined above, thus concluding that C is flow equivalent to A.

3[rd] step. *Assume A is such that $Out(I) \geq 3$ for some vertex I and $Out(J) \geq 2$ for every vertex J. Then there exists a continuous function $\alpha: \Sigma_A \to \mathbb{Z}$ with the following properties:*

(i) *$\alpha \circ g = \alpha$ for every $g \in G$;*

(ii) *there exists $t \geq 1$ such that $\alpha(x)$ depends only on $x_0, \ldots, x_{t-1}$ for every x;*

(iii) *if $b_0 \ldots b_{t-2}$ is a path in A then, denoting by $\alpha(b_0 \ldots b_{t-2}\, b_{t-1})$ the (constant) value of α on the cylinder $_0[b_0 \ldots b_{t-2}\, b_{t-1}]$, we have*

$$\sum_{b_{t-1}} 1/2^{\alpha(b_0 \cdots b_{t-2}\, b_{t-1})} = 1;$$

(iv) *there are periodic points x and y which have the same minimum period p and such that*

$$\sum_{i=0}^{p-1} \alpha(\sigma^i x) = \sum_{i=0}^{p-1} \alpha(\sigma^i y) + 1.$$

Remark. The sum in (iii) is over all the edges b_{t-1} such that $b_0 \ldots b_{t-2} b_{t-1}$ is a path in A; if $t = 1$ then instead of paths $b_0 \ldots b_{t-2}$ we fix vertices I and the sum is over the edges b_{t-1} with initial vertex I.

We first construct two cycles u and v with equal lengths which will give rise to the periodic points x and y. For our construction purposes, these cycles will have the following properties: u and gu (respectively v and gv) do not overlap if $g \neq id$; u and v do not overlap themselves except trivially; u and hv do not overlap for every $h \in G$; both u and v include the vertex I with $Out(I) = 3$.

Consider this vertex I, which by going to higher blocks we may assume to have at least two edges *into* it, and a simple cycle $e_1 \ldots e_r$ based at I. Let $I_0 = I$ and I_i be the final vertex of e_i. Let e be one edge with final vertex I which is distinct from e_r, and let f_1 and f_1 be two distinct edges emanating from I. Let K be the initial vertex of e. Going to higher blocks if necessary, we first may assume the vertex K is distinct from the vertices I_i $(0 \leq i < r)$, and then that it is even distinct from the vertices gI_i $(g \in G, 0 \leq i < r)$.

Now choose paths of the same length connecting I to K and beginning respectivelly with the edges f_1 and $\tilde{f}_1$. Let them be $f_1 \ldots f_s$ and $\tilde{f}_1 \ldots \tilde{f}_s$. Let $m \geq 1$ be so large that $m\, r \geq s - 2$. We define

$$u = e\,(e_1 \ldots e_r)^m\, f_1 \ldots f_s,$$

$$v = e\,(e_1 \ldots e_r)^m\, \tilde{f}_1 \ldots \tilde{f}_s.$$

The verification that u and v satisfy the above conditions is straightforward. We exemplify it by proving that u and gu do not overlap if $g \neq id$. What we have to check is that if we write $u = \bar{e}_0\, \bar{e}_1 \ldots \bar{e}_{t-1}$ $(t = mr + s + 1)$ then we cannot have $e_{i+j} = ge_j$ for some $0 \leq i \leq t\text{-}1$ and all $0 \leq j \leq t - 1 - i$. For assume such an i exists. To begin with i must be greater than 0, for $g \neq id$; the edge $g\bar{e}_0$ begins with the vertex gK, which does not occur in $e_1 \ldots e_r$, and this implies $i > mr + 1$; and then, in the equality $e_{t-1} = ge_{t-1-i}$, we have

$$\begin{aligned} t - 1 - i &< t - mr - 2 \\ &= s - 1 \leq mr + 1, \end{aligned}$$

which implies $\bar{e}_{t-1-i}$ is either the edge e or some edge e_j $(1 \leq j \leq r)$, and therefore the final vertex of ge_{t-1-i} is gI_j for some $0 \leq j \leq r$. This is absurd because K, the final vertex of $\bar{e}_{t-1}$, is not of this form.

We now define a function α satisfying the above conditions (i) - (iv), denoting by $\alpha(b_0\, b_1 \ldots b_{t-1})$ the value of α on the cylinder ${}_0[b_0\, b_1 \ldots b_{t-1}]$. Remember $u = \bar{e}_0\, \bar{e}_1 \ldots \bar{e}_{t-1}$ and write $v = e_0\, e_1 \ldots \hat{e}_{t-1}$. Every vertex in A has at least two outgoing edges and therefore condition (iii) forces $\alpha \geq 1$. Since the final vertex of e_0 is I, which has at least three outgoing edges, we can set $\alpha(\bar{e}_2\, e_3 \ldots \bar{e}_{t-1}\, \bar{e}_0\, \bar{e}_1) = 2$; and for the other edges b with initial vertex I we choose the values $\alpha(\bar{e}_2\, \bar{e}_3 \ldots \bar{e}_{t-1}\, e_0\, b)$ so that condition (iii) is verified. For $i \neq 2$, $0 \leq i \leq t - 1$, we set $\alpha(\bar{e}_i \ldots \bar{e}_{t-1}\, \bar{e}_0 \ldots \bar{e}_{i-1}) = 1$; and we set, for all $0 \leq i \leq t - 1$,

$\alpha(\dot{e}_i \dots \dot{e}_{t-1}\, \dot{e}_0 \dots \dot{e}_{i-1}) = 1$.

These assignments can now be extended to other t-blocks $b_0\, b_1 \dots b_{t-1}$ by using the G-action and condition (i); the non-overlapping properties of the blocks gu and hv ($g,h \in G$) ensure this can be done without conflict. The values of α on the other t-blocks are quite arbitrary, so long that we respect conditions (i) and (iii). Thus we have completed the definition of α, and we check that condition (iv) is satisfied by the periodic points $x = (u)^{\infty}$ and $y = (v)^{\infty}$.

To make the rest of the argument easier to follow we now assume $t = 1$, so that $\alpha(x)$ depends only on the edge x_0. This can be achieved by passing to the (t-1)-block system and replacing α with $\alpha \circ \sigma^{-(t-1)}$. Now consider the 0-1 matrix B defined as follows: the states are the pairs (b, j) with $1 \leq j \leq \alpha(b)$; and we allow only the transitions from (b, j) to (b, j+1) (if $j < \alpha(b)$) and from (b, α(b)) to (c, 1) (if the initial vertex of c is the final vertex of b). By 10.5 (Σ_B, G) is flow equivalent to (Σ_A, G), where the G-action on Σ_B is defined by $g(b, j) = (gb, j)$. This is a good definition because α is invariant under the G-action.

We now show B is aperiodic and has spectral radius 2. We define a stochastic matrix P with $P^0 = B$ as follows:

(1) $\quad P((b, j), (b, j+1)) = 1$ if $j < \alpha(b)$;

(2) $\quad P((b,\alpha(b)), (c,1)) = 1/2^{\alpha(c)}$ if $B((b,\alpha(b)), (c,1)) = 1$.

It is condition (iii) above that ensures P is a stochastic matrix. Now consider a simple cycle $w = b_0\, b_1 \dots b_{k-1}$ in A: this lifts to a simple cycle $\bar{w}$ in B of length $l(\bar{w}) = \alpha(b_0) + \dots + \alpha(b_{k-1})$. From (2) it is clear that $wt_P(\bar{w}) = 1/2^{l(\bar{w})}$. By 4.2 this implies m_P is the measure of maximal entropy on Σ_B, and therefore its Γ-group is generated by β^d, where β is the spectral radius and d is the period of B. Since Γ_P is clearly generated by 2, to conclude the proof we only have to show B is aperiodic. This is the work of the cycles u and v we have defined above: they lift respectively to the cycles $\bar{u}$ and $\bar{v}$ in B, and α was defined so that $l(\bar{u}) = l(\bar{v}) + 1$.□

Remarks.

1- The construction of the cycles u and v above was inspired by the similar construction in J. Ashley's paper [As].

2- Proposition 11.2 generalizes the corresponding result without group actions, which appears implicitly in [B3].

Final remarks

12. *Williams' conjecture and stochastic flow equivalence.*

At the end of section 9 of his survey article [P2], Bill Parry raises the problem of deciding when two irreducible stochastic matrices are stochastically flow equivalent. For matrices whose Γ-group is cyclic he asks whether some known invariants (such as the Γ-group, the winding numbers group W(P) and the analogues of the Parry-Sullivan invariant and the Bowen-Franks' group) form a complete set of invariants for stochastic flow equivalence. In this note we show that a positive answer to this question would settle Williams' conjecture for irreducible integer matrices affirmatively.

We briefly recall some definitions. Let A and B be square 0-1 matrices. A and B are called *shift equivalent* if there are non-negative integer matrices R, S and an integer $m \geq 1$ such that the following equalities hold:

$$A R = R B, \quad S A = B S,$$

$$R S = A^m, \quad S R = B^m.$$

Williams [W] showed that if Σ_A and Σ_B are topologically conjugate then A and B are shift equivalent; his conjecture (which remains undecided when A and B are irreducible) is that the converse also holds. There is also the notion of stochastic shift equivalence for stochastic matrices (see [PT1], [PT2] or [P2]).

Throughout this section we fix two irreducible 0-1 matrices A and B, and let P and Q be respectively the stochastic matrices which correspond to the measures of maximal entropy on Σ_A and Σ_B (recall 2.3). The proof of the following lemma is trivial.

12.1. Lemma. *If A and B are shift equivalent then P and Q are stochastically shift equivalent.*

Now assume A and B are shift equivalent. Then P and Q are (stochastically) shift equivalent and all the above-mentioned invariants for stochastic flow equivalence are the same for P and Q. This is either well-known or the proofs (in case some of them do not appear explicitly in the references we give) are straightforward. *If this set of invariants was complete* then P and Q would be stochastically flow equivalent, and our next proposition would imply that Σ_A and Σ_B are conjugate, thus proving Williams' conjecture.

12.2. Proposition. *If P and Q are stochastically flow equivalent and Σ_A and Σ_B have the same topological entropy then Σ_A and Σ_B are topologically conjugate.*

Proof. By the proof of 3.1, if P and Q are stochastically flow equivalent then there are a locally constant function f: $\Sigma_A \to \mathbf{R}^+$ and a topological conjugacy π: $\Sigma_A^f \to \Sigma_B^1$ such that $m_Q^1 = m_P^f \circ \pi^{-1}$. The measure of maximal entropy on Σ_B^1 is precisely m_Q^1, and therefore m_P^f is the measure of maximal entropy on Σ_A^f. From Proposition 6.1 in [PP] we deduce that there exists a constant $h > 0$ such that m_P is the equilibrium state of the function $-hf$, and by Proposition 3.6 [PP] we conclude that $-hf$ is cohomologous to a constant. Therefore f is also cohomologous to a constant $r > 0$, and we have a topological conjugacy π: $\Sigma_A^r \to \Sigma_B^1$. By the result in [A], if h_A is the topological entropy of Σ_A then the topological entropy of Σ_A^r is equal to $(1/r)h_A$, and (since $h_A = h_B$) this implies $r = 1$. Thus we have proved the existence of a topological conjugacy π: $\Sigma_A^1 \to \Sigma_B^1$, and from here (as in the proof of 3.1) the existence of a topological conjugacy π_0: $\Sigma_A \to \Sigma_B$ follows easily.□

References

[A] Abramov, L.M.: On the entropy of a flow. Dokl. Akad. Nauk. SSSR 128, 647-650 (1959)

[AKM] Adler, R.L., Kitchens, B., Marcus, B.: Finite group actions on shifts of finite type. Erg. Th. and Dyn. Syst. 5, 1-25 (1985)

[AM] Adler, R.L., Marcus, B.: Topological entropy and equivalence of dynamical systems. Memoirs A.M.S. 219 (1979)

[As] Ashley, J.: Bounded to one factors of an aperiodic shift of finite type are one to one almost everywhere factors also. Erg. Th. and Dyn. Syst. 10, 615-625 (1990)

[B1] Bowen, R.: Symbolic dynamics for hyperbolic flows. Amer. J. of Math. 95, 429-460 (1973)

[B2] Bowen, R.: Equilibrium states and the ergodic theory of Anosov diffeomorphisms. (Lect. Notes Math., vol. 470) Berlin Heidelberg New York: Springer 1975

[B3] Boyle, M.: Almost flow equivalence for hyperbolic basic sets. To appear in Topology.

[F] Franks, J.: Flow equivalence of subshifts of finite type. Erg. Th. and Dyn. Syst. 4, 53-66 (1984)

[K1] Krieger, W.: On the finitary isomorphisms of Markov shifts that have finite-expected coding time. Z. Wahr. 65, 323-328 (1977)

[K2] Krieger, W.: On the subsystems of topological Markov chains. Erg. Th. and Dyn. Syst. 2, 195-202 (1982)

[L] Lind, D.: The entropies of topological Markov shifts and a related class of algebraic integers. Erg. Th. and Dyn. Syst. 6, 571-582 (1986)

[MT] Marcus, B., Tuncel, S.: The weight-per-symbol polytope and scaffolds of invariants associated with Markov chains. Erg. Th. and Dyn. Syst. 11, 129-180 (1991)

[N] Newman, M.: Integral Matrices. New York: Academic Press (1972)

[P] Parry, W.: Intrinsic Markov chains. Trans. A.M.S. 112, 55-65 (1964)

[P2] Parry, W.: Notes on coding problems for finite state processes. Bull. London Math. Soc. 23, 1-33 (1991)

[PP] Parry, W., Pollicott, M.: Zeta functions and the periodic orbit structure of hyperbolic dynamics. Astérisque 187-188 (1990)

[PSc] Parry, W., Schmidt, K.: Natural coefficients and invariants for Markov chains. Invent. Math. 76, 15-32 (1984)

[PSu] Parry, W., Sullivan, D.: A topological invariant for flows on one-dimensional spaces. Topology 14, 297-299 (1975)

[PT1] Parry, W., Tuncel, S.: On the stochastic and topological structure of Markov chains. Bull. L.M.S. 14, 16-27 (1982)

[PT2] Parry, W., Tuncel, S.: Classification problems in Ergodic Theory. (L.M.S. Lect. Notes, vol. 67) Cambridge: Cambridge University Press 1982

[Pe] Petersen, K.: Chains, entropy, coding. Erg. Th. and Dyn. Syst. 6, 415-448 (1986)

[R] Ruelle, D.: Thermodynamic formalism. (Encyclopedia of Mathematics and its Applications, vol. 5) Reading, Mass: Addison-Wesley 1978

[W] Williams, R.F.: Classification of sub-shifts of finite type. Ann. of Math. 98, 120-153 (1973). Errata. Ann. of Math. 99, 380-381 (1974)